FUNDAMENTAL PRINCIPLES OF

Physical Chemistry

FUNDAMENTAL PRINCIPLES OF

Physical Chemistry

CARL F. PRUTTON, Ph.D.

Professor of Chemistry and Chemical Engineering
Case School of Applied Science

SAMUEL H. MARON, Ph.D.

Associate Professor of Physical Chemistry
Case School of Applied Science

THE MACMILLAN COMPANY · NEW YORK

1944

Preface

The authors' aim in writing this text on elementary physical chemistry is to place in the hands of teachers and students a book which covers the fundamental principles of the subject in a thorough, sound, up-to-date, and clear manner.

In deciding what constitutes the fundamental principles the authors were continually guided by the needs of the chemist and chemical engineer for a sound grounding in physical chemistry. Although physical chemistry is offered to various students with various purposes in mind, the fact remains that those who expect to be engaged in any branch of chemical or related work must be conversant with the principles of this subject, and they must be able to use and apply these principles effectively and correctly. In order to do this they must be exposed to a basic training sufficiently complete to permit them to understand the subject, not only in a general and qualitative way, but also in its more intimate experimental and quantitative aspects. To achieve such mastery the authors feel that use of mathematics and some thermodynamics is absolutely essential. For this reason the necessary calculus is employed here without any apology, and the elements of thermodynamics are introduced early and are used throughout the book as an integral part of the subject.

Any physical chemist is well aware that the development of physical chemistry did not stop at a certain period nor did it start there. Rather, like any other rapidly moving scientific field, physical chemistry has undergone natural changes and refinements which can neither be ignored nor magnified out of their true proportions. In writing this book a very conscious effort was made to present each subject in a balanced form. By doing this the authors feel that they have been able to outline physical chemistry as it actually is, and not as a progression of ideal gases, ideal solutions, and idealized electrochemistry on the one hand, or quantum theory and mechanics on the other. Further, at no time is there any timidity exhibited in stating that a particular principle has its limitations, or that a particular equation is not always valid. It is true that by such a procedure it may not be possible to create in the student an impression that physical chemistry as we know it today is infallible and perfect.

However, compensation is found in the fact that the gap which frequently exists between a student's training on the undergraduate and graduate levels is narrowed, and the resentment and futility which assail a student when he is forced to unlearn what was presented to him as the gospel truth are avoided.

Although the authors claim no radical innovations in presenting the subject matter, they believe that there are a number of desirably distinctive features in the book. The use of thermodynamics for clarity and rigidity of treatment has already been alluded to. In discussing any particular subject an effort is made to present the requisite argument in some detail, to bring the subject as much as possible up-to-date, and to reduce to a minimum the introduction of equations and principles merely on the authors' say so. For this reason the second chapter, immediately after gases, is devoted to an elementary treatment of the first and second laws of thermodynamics in order that the Clausius-Clapeyron and Clapeyron equations may be introduced in the next two chapters on a rational basis. In general each subject is presented as completely as is deemed necessary, and with a critical attitude. Throughout the book the effort may be discernible to systematize and organize the material in a logical manner, and to show the connection and relation of various topics.

In particular the authors wish to direct attention to their treatment of chemical equilibrium and electrochemistry. Chemical equilibrium is introduced from both the kinetic and thermodynamic points of view. From the latter approach the student is made aware of the approximations involved in ordinary formulations, and he is put in a position to understand what is required to make the treatment exact. Heterogeneous equilibria are treated from the standpoints of the law of mass action, the distribution law, and the phase rule. Especially is the phase rule presented in considerable detail, for it is the authors' belief that all too frequently is this highly important subject sadly neglected.

The use of thermodynamics permits the introduction of the activity concept without any arbitrariness, and allows ionic equilibria and electrochemistry to be handled in a satisfactorily rigid manner. Sufficient indications are given to enable a student to decide where activities may be approximated by concentrations without serious error, and where activity coefficients must be used. The Debye-Hückel theory of electrolytes is accorded due attention here. The authors feel that their presentation of both electromotive force and polarization is more complete and exact than usual, and that the subjects are discussed in sufficient detail to permit a student to orient himself in electrochemistry as it is actually encountered in the laboratory and literature. Further, both these subjects are treated from a unified point of view and in terms of the generally prevalent notation of G. N. Lewis. In this connection it may be pointed out that a simple

scheme is included for arriving at the signs of potentials and at electrode and cell reactions. This scheme should help greatly in overcoming a good many of the difficulties experienced by students.

Throughout the book considerable attention is devoted to the experimental aspects of physical chemistry. This is done expressly, because it is felt that a student can understand better what he is dealing with when he has some idea how the quantities involved are obtained. At the end of each chapter are given bibliographies, to which the student may turn for further details, and extended lists of exercises to test his familiarity with the subject matter and to develop in him facility in handling equations and data.

The book as a whole is intended primarily for a full year course in physical chemistry for students of chemistry and chemical engineering. By judicious selection of various portions of the contents, this book should be readily adaptable to any one-semester course for students in other branches of science, such as physicists, metallurgists, biologists, and biochemists. What is to be omitted and what is not is largely a matter of circumstances and needs, and can be left safely to the discretion and decision of each individual teacher.

The authors gladly acknowledge the kindness shown by Professor David Turnbull in reading the entire manuscript, and in offering helpful suggestions and criticisms.

Table of Contents

ix

FUNDAMENTAL PRINCIPLES OF

Physical Chemistry

Introduction

Science and the Scientific Method. Science is organized and systematized knowledge relating to our physical world. This knowledge did not spring into being full blown, but has been accumulated painstakingly through the efforts of many researchers and observers. In its inception this cumulative process was quite simple. It involved merely the observation of phenomena as they occur in nature and their faithful recording. As the facts and observations multiplied, regularities were sought and discovered in them which were then formulated into *laws*. Each law was capable of embracing a number of facts and of summarizing them in succinct form.

However, natural laws do not constitute an interpretation of nature, but rather a description. To supply the reason for the operation of a law scientists began to propose purely suggestive explanations, or *hypotheses*, in terms of which the operation of the law could be accounted. From these hypotheses logical predictions were then derived and compared with the observed phenomena. If the two tallied fairly well, the hypothesis was accepted, provisionally at least, and became a *theory*. On the other hand, if the logical deductions of a hypothesis did not agree with experimental facts, the hypothesis was discarded to await a more satisfactory explanation.

At present we do not rely on purely fortuitous observation for our information. With the state of scientific knowledge as advanced as it is, experiments are carefully planned and conducted to yield the specific data sought. It is thus possible to arrive at desired facts more rapidly and efficiently. Again, with planned research hypotheses and theories may be subjected to deliberate test by arranging experiments designed to answer directly the specific points in question. By such means faulty concepts can be eliminated, incomplete theories refined, and new principles discovered.

This modus operandi of science is called the *scientific method*. It will be noticed that in its operation the scientific method involves essentially four stages, namely: (1) the accumulation of facts; (2) the organization of facts into laws; (3) the postulation of hypotheses to account for the facts and the laws; and (4) comparison of the hypothetical deductions with the experimental results. Whenever possible, facts and their correlations, as well as their explanations, are expressed in mathematical

terms. It is this precision of language more than any other single factor which differentiates the physical sciences of physics and chemistry from the more purely descriptive sciences of biology or medicine.

The intimate combination of experiment and theory embodied in the scientific method has proved very fruitful, and has led to the development of our present highly advanced state of science and technology. It must be emphasized, however, that the function of theory and hypothesis in this advance has not been merely to explain what is already known. Were this the only contribution of theory, speculation would have been more of interest than value. The real function of theory and speculation lies much more intrinsically in its ability to define the experimental variables, and in its ability to foretell phenomena and effects that are as yet unknown. When thus used, theory and hypothesis may serve not only as a powerful guide in the interpretation of phenomena, but also as an effective tool for the advancement of our knowledge of the physical world and its control for our benefit.

Physical Chemistry. The branch of chemistry which concerns itself with the study of the physical properties and structure of matter, with the laws of chemical interaction, and with the theories governing these is called *physical chemistry*. The purpose of physical chemistry is, first of all, to collect the appropriate data required to define the properties of gases, liquids, solids, solutions, and colloidal dispersions, to systematize them into laws, and to give them a theoretical foundation. Next physical chemistry is interested in establishing the energy relations obtaining in physical and chemical transformations, to ascertain their extent and the speed with which they take place, and to define quantitatively the controlling factors. In this connection must be considered not only the more common variables of temperature, pressure, and concentration, but also the effects of the intimate interaction of matter with electricity and light. Finally matter itself must be examined to ascertain its nature and structure. This is necessary in order that we may be able to arrive at a basic understanding of physical and chemical behavior in terms of the properties of the fundamental constituents of matter itself.

To accomplish its purposes physical chemistry must rely to a large degree on experiment. Experimental methods and techniques play thus a very important role. The subject also draws generously on the laws and methods of physics and mathematics. In fact, in many respects physical chemistry may be looked upon as the field where physics and mathematics are applied extensively to the study and solution of problems of prime chemical interest. With the appropriate data at hand, physical chemistry then proceeds to its correlational and theoretical goal through two general modes of attack, namely, the *thermodynamic*

and the *kinetic.* In the thermodynamic approach the fundamental laws of thermodynamics are utilized to yield deductions based on the energy relations connecting the initial and final stages of a process. By circumventing the steps intervening between the start and end of a process, thermodynamics enables us to arrive at many valuable deductions without our knowing all the intimate details of the intermediate stages. Consequently, although this approach is able to tell us what can happen, and to what extent, it is unable, by its very nature, to give us information on *how,* or *how rapidly,* a change will actually occur. On the other hand, the kinetic approach requires for its operation an intimate and detailed "picture" of the process. From the mechanism postulated may be deduced then the law for the overall process and its various stages. Evidently the kinetic approach to a problem is more explanatory in character, but unfortunately it is generally more complicated and difficult to apply. These two modes of attack will be illustrated at various stages in the text. From the examples given there the student will be able to differentiate more clearly between them and come to appreciate their respective powers and utilities.

History of Physical Chemistry. The roots of physical chemistry lie in the fields of both chemistry and physics. At first these two branches of science developed more or less independently. However, in the nineteenth century it was found that the discoveries in physics had important bearing on and application to chemistry, and hence a need arose for a more or less distinct field dealing primarily with the application of physical laws to chemical phenomena. This need finally impelled Wilhelm Ostwald, van't Hoff, and Arrhenius to organize and systematize the subject matter generally included now under the head of physical chemistry, and led them in 1881 to found the *Zeitschrift für physikalische Chemie.* The inception of physical chemistry as a formal branch of chemical science may be dated from the appearance of this journal.

Stimulated by this publication, and fostered by the contributions of the men mentioned, physical chemistry entered a period of very rapid growth. Aiding this progress were not only advances in chemistry, but also the remarkable series of discoveries in physics which started with the discovery of the electron, and which include the discovery of x rays and radioactivity, the establishment of the quantum theory, and the unfolding of our understanding of subatomic phenomena. Thanks to these contributions, physical chemistry has developed in the past 60 years or so to a position of importance and utility not only to chemistry but to other sciences as well.

Importance of Physical Chemistry. Since physical chemistry deals with the principles and theories of chemistry, it goes without say-

ing that any student or practitioner of this science must be familiar with
the fundamental principles of physical chemistry in order to understand
his own subject. The same applies also to the chemical engineer. The
essential difference between a chemist and a chemical engineer is that
whereas the former conducts his reactions and operations on a small
scale, the chemical engineer carries them out in large commercial units.
To transfer an operation from the laboratory to a plant the chemical
engineer must of course be able to apply engineering and economic prin-
ciples. However, at the same time he must understand also the funda-
mentally chemical nature of the processes he is dealing with, and for that
he needs physical chemistry. As a matter of fact, chemical engineering
has frequently been described as applied physical chemistry. Viewed in
this light, many of the aspects of chemical engineering fall within the
realm of physical chemistry and can be handled in terms of well-estab-
lished and familiar physicochemical laws. On the other hand, any
attempt to consider chemical engineering as a purely empirical pursuit
robs it of the attributes of a science and translates it back again into
an art.

What has been said about the importance of physical chemistry to
the chemist and chemical engineer applies equally well to the metallur-
gist and metallurgical engineer. The latter two perform essentially the
same functions as the two former, except that their attention is confined
primarily to metals. From this point of view the prominent position
of physical chemistry, whether under this or other titles, in these sub-
jects becomes clear, and accounts for the valuable contributions made
to these fields by the application of physicochemical principles.

Finally, physical chemistry finds application also in physics, geology,
and in the various ramifications of the biological sciences. To appreciate
the extent of its utility it is only necessary to compare a book on chemical
physics, geology, or biochemistry with one on physical chemistry. From
such a comparison it becomes quite evident why physical chemistry is
often included in curricula in these subjects, and why it can be applied
with effect in these sciences.

Scope of Book. The present text starts with a discussion of the
behavior of gases, liquids, solids, solutions, and colloids. This discussion
embraces the first eight chapters, except for Chapter II, where certain
necessary elementary principles of thermodynamics are introduced.
These principles of thermodynamics are extended then in Chapters IX,
X, and XI, and are applied to a consideration of chemical equilibria in
Chapters XII, XIII, and XIV. Beginning with Chapter XV and extend-
ing through Chapter XVIII is given an exposition of the principles of
electrochemistry, i.e., the interaction of matter and electricity. This is
followed by a consideration of the rates of homogeneous reactions,

Chapter XIX, and heterogeneous reactions, Chapter XX. Finally, in Chapters XXI and XXII accounts are given of our knowledge of atomic and molecular structure, while in the last chapter the information of atomic and molecular structure is applied to correlation of physical properties with chemical constitution.

CHAPTER I

Gases

The States of Aggregation. All matter as we know it exists in one of three states of aggregation, solid, liquid, or gaseous. A solid may be defined as a body possessing both definite volume and definite shape at a given temperature and pressure. Under an applied force a solid may be distorted and suffer a change in volume, but, provided the elastic limit has not been exceeded, removal of the force will restore the body to its original condition. A liquid, on the other hand, has a definite volume but no definite shape, while a gas has neither definite shape nor volume. Liquids and gases are both termed *fluids*, and both offer no resistance to shape deformation. Inertia and viscosity may resist changes in shape, but these forces have no tendency to restore the fluid to its original form. A liquid, insofar as it fills the container, will always adopt the shape of the container in which it is placed, but will retain its definite volume, while a gas will always fill completely any container in which it may be confined.

The distinctions among the three states of matter are not always as clear cut as the above definitions would imply. For example, a liquid at the critical point is indistinguishable from its vapor. Again, such substances as glass or asphalt, although exhibiting many of the properties of a solid, will, under certain conditions of temperature, become plastic and exhibit properties not ascribed to pure solids. For this reason such substances are usually considered to be supercooled liquids with very high viscosity.

The particular state of aggregation of a substance is determined by the temperature and pressure under which it exists. However, within certain limits of temperature and pressure a substance may exist in more than one state at the same time. In fact, under special conditions a substance may exist in all three states simultaneously. Thus at 4.57 mm Hg pressure and at 0.0099° C, ice, water, and water vapor may all be present simultaneously, and all be stable. This subject of simultaneous existence in more than one state will be discussed more completely in subsequent chapters.

Nature of Gases. A gas may be regarded as consisting of molecules traveling in straight lines at random and at high rates of speed within

the containing space, and colliding frequently with other molecules or the walls of the container. The force exerted per unit area on the walls of the container by the colliding molecules is known as the *pressure* — a force present at all times and distributed uniformly over the entire surface. The fact that small molecules produce a considerable bombarding force upon container walls suggests that the number of collisions with the walls must be large and that the molecules must be moving with high velocities.

The space occupied by the molecules themselves within a gaseous volume is a small fraction of the total volume of the gas under ordinary conditions of temperature and pressure. Thus, if all the air in a room 20 by 10 by 10 ft were liquefied, the volume of the liquid would be approximately 2.4 cu ft, or about 0.1 per cent of the volume of the room, and yet the molecules would not be touching each other. Hence we may conclude that molecules generally are separated from each other by distances which are large compared to molecular diameters, and that within a gas the space actually occupied by molecules is very small, most of the volume being "free" space. This accounts for the much lower densities of gases as compared to liquids and solids.

Also, this large amount of "free" space within a gas makes compression of the gas fairly easy. The compression process merely reduces the large "free" space and, by reducing the average distance between the molecules, brings them closer together. When there is no attraction between the molecules, the decrease in "free" space on compression is equal to the observed decrease in the total volume of the gas. Similarly, on expansion the average distance between molecules is increased, and thereby also the "free" space of the gas. In any case the random motion of the molecules will give the effect of completely filling any containing vessel in which the gas is placed.

In terms of the structure of a gas outlined above it is easy to understand why gases interdiffuse or mix. Two different gases such as nitrogen and oxygen, or any number of nonreactive gases, when placed in a container will by their motion mix with one another very quickly regardless of density. This mixture of gases will in many respects behave like a single gas, and the molecules of the various gases will collide with each other regardless of similarity or dissimilarity. Further, the total pressure exerted by the mixture will be determined by the total number of collisions between the molecules of all kinds and the walls of the container, a pressure to which each particular kind of molecule contributes its share.

Ideal and Real Gases. In discussing their laws of behavior, gases will frequently be referred to as being either *ideal*, or *nonideal* or *real*. An ideal gas is one that obeys certain laws which will be presented shortly, while a real gas is one that obeys these laws only at low pressures.

The deviations from the ideal laws are due in general to two factors of which the ideal laws take no account, namely, the volume actually occupied by the molecules themselves, and the attractive forces existing between the molecules.

An ideal gas is one in which the volume occupied by the molecules themselves is negligible compared to the total volume at all pressures and temperatures, and one in which the intermolecular attraction is extremely small under all conditions. In a nonideal or real gas both of these quantities are appreciable, the magnitude of each depending on the nature, the temperature, and the pressure of the gas. We can easily see that an ideal gas must be a hypothetical gas, as all actual gases must contain molecules which occupy a definite volume and exert attractions between each other. However, very often the influence of these factors becomes negligible, and the gas then may be considered to be ideal. We shall find that the latter condition will obtain in particular at low pressures and relatively high temperatures, conditions under which the "free" space within the gas is large and the attractive forces between molecules small.

Empirical Generalizations of Ideal Gas Behavior. Through the study of gases there have been evolved certain laws or generalizations which are always the starting point in any discussion of gas behavior. These are: (1) Boyle's law, (2) Charles's or Gay-Lussac's law, (3) Dalton's law of partial pressures, and (4) Graham's law of diffusion. Another generalization is Avogadro's principle, but this will be considered later. These laws will now be discussed in some detail.

Boyle's Law. In 1662 Robert Boyle reported to the Royal Society of England the results of his studies on the relation between the volume and pressure of a gas at constant temperature. Boyle confined within a graduated tube a quantity of gas, and then measured the volume of that gas under different applied pressures. He found that the volume decreased with increasing pressure and that, within the limits of his experimental accuracy, *the volume of any definite quantity of gas at constant temperature varied inversely as the pressure on the gas.* This highly important generalization is known as *Boyle's law.* Expressed mathematically this law states that at *constant temperature* $V \propto 1/P$

or that
$$V = \frac{C}{P}$$

where V is the volume and P the pressure of the gas, while C is a proportionality constant dependent on the temperature and quantity of gas involved. On rearrangement this equation becomes

$$PV = C \tag{1}$$

from which it follows that if in a certain state the pressure and volume of the gas are P_1 and V_1, while in another state they are P_2 and V_2, then at constant temperature

$$P_1V_1 = C = P_2V_2$$

and

$$\frac{P_1}{P_2} = \frac{V_2}{V_1} \tag{2}$$

If the pressure of a gas is plotted against the volume in accordance with equation (1), we obtain a family of curves such as that shown in Fig. 1. Each curve is a hyperbola with a different value of C. Since for
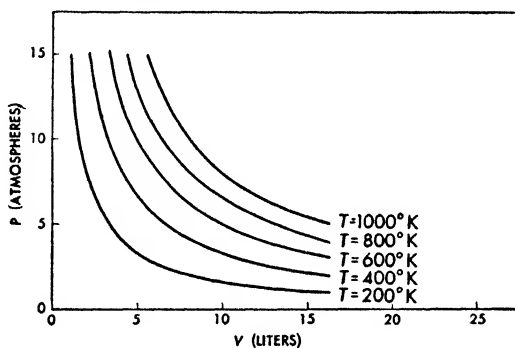

Fig. 1. Isothermal Plot of P vs. V according to Boyle's Law (One Mole of Gas)

a given weight of gas C varies only with temperature, each curve corresponds to a different fixed temperature and is known as an *isotherm* (constant temperature plot). The higher curves correspond to the higher temperatures.

The Charles or Gay-Lussac Law. Charles in 1787 observed that the gases hydrogen, air, carbon dioxide, and oxygen expanded an equal amount upon being heated from 0 to 80° C at constant pressure. However, it was Gay-Lussac in 1802 who first made a quantitative study of the expansion of gases on heating. He found that for *all* gases the increase in volume for each degree centigrade rise in temperature was equal approximately to 1/273 of the volume of the gas at 0° C. A more precise value of this fraction is 1/273.15. If we designate by V_0 the volume of a gas at 0° C and by V its volume at any temperature $t°$ C, then in terms of Gay-Lussac's finding V may be written as

$$V = V_0 + \frac{t}{273.15} V_0$$

$$= V_0\left(1 + \frac{t}{273.15}\right)$$

$$= V_0\left(\frac{273.15 + t}{273.15}\right) \tag{3}$$

We may define now a new temperature scale such that any temperature t on it will be given by $T = 273.15 + t$, and $0°$ C by $T_0 = 273.15$. Then equation (3) becomes simply,

$$\frac{V}{V_0} = \frac{T}{T_0}$$

or generally

$$\frac{V_2}{V_1} = \frac{T_2}{T_1} \qquad (4)$$

This new temperature scale, designated as the absolute, Kelvin, or thermodynamic scale of temperature, is of fundamental importance in all science. In terms of this temperature scale, equation (4) tells us that *the volume of a definite quantity of gas at constant pressure is directly proportional to the absolute temperature*, or that

$$V = kT \qquad (5)$$

where k is a proportionality constant determined by the pressure, the amount of gas, and the units of V. The above statement and equation (5) are expressions of *Charles's or Gay-Lussac's law of volumes.*
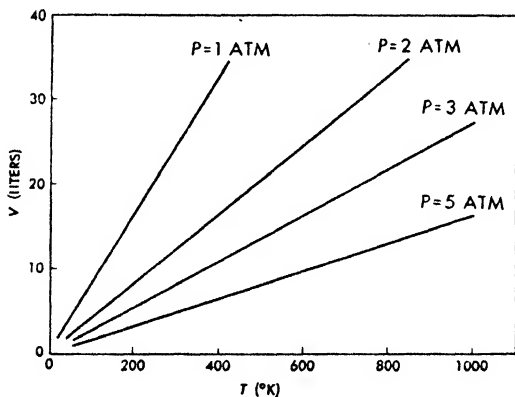

Fig. 2. Isobaric Plot of V vs. T according to Charles's Law (One Mole of Gas)

According to equation (5) the volume of a gas should be a straight line function of the absolute temperature at any constant pressure. Such a plot of V vs. T at selected pressures is shown in Fig. 2. Since for a given amount of gas k will have different values at different pressures, we obtain a series of straight lines, one for each constant pressure. Each constant pressure line is called an *isobar*. For every isobar the slope is the greater the lower the pressure.

Equation (5) suggests also that if we were to cool a gas to $0°$ K ($- 273°$ C), its volume would become zero. However, no such phenomenon is ever encountered, for usually long before $0°$ K is approached a

gas liquefies or solidifies. Again, as will be shown below, under such drastic conditions the equation itself cannot be considered to hold.

The Combined Gas Law. The two laws discussed give the separate variation of the volume of a gas with pressure and with temperature. To obtain the simultaneous variation of the volume with temperature and pressure, we proceed as follows. Consider a quantity of gas at P_1, V_1, and T_1, and suppose that it is desired to obtain the volume of the gas, V_2, at P_2 and T_2. First of all let us compress (or expand) the gas from P_1 to P_2 at constant temperature T_1. The resulting volume V_x then will be, according to Boyle's law,

$$\frac{V_x}{V_1} = \frac{P_1}{P_2}$$

$$V_x = \frac{V_1 P_1}{P_2} \tag{6}$$

If the gas at V_x, P_2, and T_1 is heated now at constant pressure P_2 from T_1 to T_2, the final state at P_2 and T_2 will have the volume V_2 given by Charles's law, namely,

$$\frac{V_2}{V_x} = \frac{T_2}{T_1}$$

$$V_2 = \frac{V_x T_2}{T_1}$$

Substituting into this relation the value of V_x from equation (6), V_2 becomes

$$V_2 = \frac{V_x T_2}{T_1} = \frac{P_1 V_1 T_2}{P_2 T_1}$$

and on rearranging terms we see that

$$\frac{P_1 V_1}{T_1} = \frac{P_2 V_2}{T_2} = \text{constant} = K \tag{7}$$

i.e., the ratio PV/T for any given state of a gas is a constant. Consequently we may drop the subscripts and write for any gas which obeys Boyle's and Charles's laws

$$PV = KT \tag{8}$$

Equation (8) is known as the *combined gas law*, because it represents a combination of Boyle's and Charles's laws. It gives the complete relationship between the pressure, volume, and temperature of any gas as soon as the constant K is evaluated. That Boyle's and Charles's laws are merely special cases of equation (8) is easily shown. When T is con-

stant, equation (8) reduces to PV = constant, or Boyle's law. Again, when P is constant, equation (8) becomes

$$V = \frac{K}{P} T = kT$$

or Charles's law.

The Gas Constant. The numerical value of the constant K in equation (8) is determined by the amount of gas involved and the units in which P and V are expressed; but it is *totally independent of the nature of the gas.* Equation (8) shows that for any given pressure and temperature an increase in the quantity of gas increases the volume, and thereby also correspondingly the magnitude of K. In other words, K is directly proportional to the number of moles [1] of gas involved. For convenience this constant may be replaced, therefore, by the expression $K = nR$, where n is the number of moles of gas occupying volume V at P and T, while R is the *gas constant per mole.* Thus expressed R becomes a *universal constant for all gases* and equation (8) takes the final form

$$PV = nRT \tag{9}$$

Equation (9) is the *ideal gas equation,* one of the most important relations in physical chemistry. It connects directly the volume, temperature, pressure, and number of moles of a gas, and permits all types of gas calculations as soon as the constant R is known. R may be found from the experimental fact that 1 mole of *any* ideal gas at standard conditions, i.e., at $0°$ C and 1 atmosphere pressure, occupies a volume 22.415 liters. If we express then the *volume in liters* and the *pressure in atmospheres,* R follows from equation (9) as

$$R = \frac{PV}{nT} = \frac{1 \times 22.415}{1 \times 273.15} = 0.08206 \text{ liter-atm/degree/mole}$$

This value of R can be used only when volume is taken in liters and pressure in atmospheres. For other combinations of units R will have other values. Thus, if the pressure be expressed in atmospheres while the volume in cubic centimeters, R becomes

$$R = \frac{1 \times 22,415}{1 \times 273.15} = 82.06 \text{ cc-atm/degree/mole}$$

Since pressure is force per unit area and volume is area times length, it immediately follows that the units of PV/nT and hence of R are:

---

[1] A *mole* is the mass of a substance in grams equal numerically to its molecular weight.

$$\frac{PV}{nT} = R = \frac{\dfrac{\text{force}}{\text{area}} \times \text{area} \times \text{length}}{\text{moles} \times \text{degrees}} = \frac{\text{force} \times \text{length}}{\text{mole} \times \text{degree}} = \frac{\text{work}}{\text{mole} \times \text{degree}}$$

Consequently R may be expressed in any set of units representing work or energy. Although in gas calculations in the metric system the units given above are the most useful, there is necessity in other types of calculations to employ R in some alternate energy units. These are usually ergs, joules, and calories.

To obtain R in ergs the pressure must be expressed in dynes per square centimeter and the volume in cubic centimeters. For the volume at standard conditions we have $V = 22{,}415$ cc. Again, a pressure of 1 atmosphere is the pressure of a column of mercury 76 cm high and 1 sq cm in cross section at $0°$ C. The total volume of such a column is thus 76 cc, and the mass 76×13.596, where the latter quantity is the density of mercury at $0°$ C. The pressure in dynes per square centimeter will be then this mass multiplied by the acceleration of gravity, 980.6 cm per second per second, or $P = 76 \times 13.596 \times 980.6$ dynes per square centimeter. Inserting these values of V and P into the expression for R, we find that

$$R = \frac{(76 \times 13.596 \times 980.6)(22{,}415)}{1 \times 273.15} = 8.315 \times 10^7 \text{ ergs/mole/degree}$$

Further, since 1 joule $= 10^7$ ergs, and 1 calorie $= 4.185$ joules, we arrive also at

$$R = 8.315 \text{ joules/degree/mole}$$
$$= \frac{8.315}{4.185} = 1.987 \text{ cal/degree/mole}$$

It should be clearly understood that, although R may be expressed in various units, for pressure-volume calculations involving gases R *must always be taken in the same units as those used for pressure and volume.* In other words, for pressure in atmospheres and volume in liters $R = 0.08206$ liter-atmosphere, while for pressure in atmospheres and volume in cubic centimeters, $R = 82.06$ cc-atmospheres. When other units are encountered it is generally simpler to convert them to those given above than to find the corresponding values of R.

In chemical engineering calculations are frequently made employing English instead of metric units, and using the pound-mole, i.e., the weight in pounds corresponding to the molecular weight, rather than the gram-mole. Further, the temperature may be in degrees Fahrenheit, $°$ F, in which case the absolute temperature is expressed in degrees

Rankine, $^\circ R = 460 + {}^\circ F$. To facilitate such calculations a summary of the values of R in various units is given in Table 1.

TABLE 1

VALUES OF R IN VARIOUS UNITS

Units of Pressure	Units of Volume	Temperature	n	R
Atmospheres	liters	$^\circ K$	gram-moles	0.08206 liter-atm/$^\circ$ K/mole
Atmospheres	cc	$^\circ K$	gram-moles	82.06 cc-atm/$^\circ$ K/mole
Dynes/cm²	cc	$^\circ K$	gram-moles	8.315×10^7 ergs/$^\circ$ K/mole
Atmospheres	cu ft	$^\circ R$	pound-moles	0.729 cu ft-atm/$^\circ$ R/mole
Pounds/sq in.	cu ft	$^\circ R$	pound-moles	10.71 cu ft-lb/in.²/$^\circ$ R/mole
Pounds/sq in.	cu in.	$^\circ R$	pound-moles	18,510 cu in.-lb/in.²/$^\circ$ R/mole
R in joules		$^\circ K$	gram-moles	8.315 joules/$^\circ$ K/gram-mole
R in calories		$^\circ K$	gram-moles	1.987 cal/$^\circ$ K/gram-mole

Calculations Involving Ideal Gas Law. The ideal gas law may be employed to find any one of the variables P, V, T, or n from any specified set of three of these. As an illustration, suppose that we want to know what will be the volume occupied by 10 grams of oxygen at 25° C and 650 mm Hg pressure. From the data we have that:

$$n = \frac{10}{32} = 0.312 \text{ mole}$$
$$T = 273.2 + 25 = 298.2^\circ \text{ K}$$
$$P = \frac{650}{760} = 0.855 \text{ atm}$$
$$R = 0.0821 \text{ liter-atm}$$

Insertion of these into equation (9) yields for the volume:

$$V = \frac{nRT}{P} = \frac{0.312 \times 0.0821 \times 298.2}{0.855}$$
$$= 8.94 \text{ liters}$$

Similarly, from appropriately specified data the other quantities involved in the ideal gas equation may be found.

Dalton's Law of Partial Pressures. It has already been pointed out that different gases introduced into the same container interdiffuse or mix rapidly and that the pressure of such a mixture is the sum total

of the contributions of the individual gases. The question now is: How much does each gas contribute to the total pressure? The answer to this question is embodied in the *law of partial pressures* which was first discovered by Cavendish in 1781 but not announced until 29 years later by Dalton. The law states that *at constant temperature the total pressure exerted by a mixture of gases in a definite volume is equal to the sum of the individual pressures which each gas would exert if it occupied the same total volume alone.* In other words, what this law says is that, in any mixture of gases in a given volume at constant temperature, the total pressure of the mixture may be considered equal to the sum of the individual pressure contributions of the various gases, namely,

$$P_{\text{total}} = P_1 + P_2 + P_3 + \cdots \tag{10}$$

The individual pressures, P_1, P_2, P_3, are termed the *partial pressures* of the respective gases. The partial pressure of each constituent may be thought of as the pressure which that constituent would exert if it were isolated in the same volume and at the same temperature as that of the mixture. In terms of the partial pressures, Dalton's law may be restated as follows: *The total pressure of a mixture of gases is equal to the sum of the partial pressures of the individual components of the mixture.*

The significance of Dalton's law and of the concept of partial pressures is best brought out by the following example. If we were to take three *1-liter* flasks filled respectively with hydrogen at 70 mm Hg pressure, carbon monoxide at 500 mm, and nitrogen at 1000 mm, all at the same temperature, and were to force all these gases into a fourth *1-liter* flask, the total pressure within the fourth flask would be

$$P = P_{\text{H}_2} + P_{\text{CO}} + P_{\text{N}_2}$$
$$= 70 + 500 + 1000$$
$$= 1570 \text{ mm Hg}$$

and the pressures of the individual gases within their 1-liter flasks would be the partial pressures of these gases in the mixture.

Consider now a gaseous mixture composed of n_1 moles of one gas, n_2 moles of another gas, and n_3 moles of still a third. Let the total volume be V and the temperature T. If the conditions of pressure and temperature are not too extreme, the ideal gas laws would be valid for each gas in the mixture, and we obtain for the respective partial pressures:

$$P_1 = \frac{n_1 R T}{V} \tag{11a}$$

$$P_2 = \frac{n_2 R T}{V} \tag{11b}$$

$$P_3 = \frac{n_3 R T}{V} \tag{11c}$$

But according to Dalton's law the total pressure is $P = P_1 + P_2 + P_3$, and hence,

$$P = \frac{n_1 RT}{V} + \frac{n_2 RT}{V} + \frac{n_3 RT}{V}$$

$$= \frac{(n_1 + n_2 + n_3)RT}{V}$$

$$= \frac{n_t RT}{V} \tag{12}$$

where $n_t = (n_1 + n_2 + n_3)$ = total number of moles of gas in the mixture. We see from equation (12), therefore, that the gas laws may be applied to mixtures as well as to pure gases, and in exactly the same way.

On division of equation (11a) by equation (12) it is found that

$$\frac{P_1}{P} = \frac{\dfrac{n_1 RT}{V}}{\dfrac{n_t RT}{V}} = \frac{n_1}{n_t}$$

and

$$P_1 = \frac{n_1}{n_t} P \tag{13a}$$

The same operations with equations (11b) and (11c) yield similarly:

$$P_2 = \frac{n_2}{n_t} P \tag{13b}$$

$$P_3 = \frac{n_3}{n_t} P \tag{13c}$$

Equations such as (13) are very important in chemical and chemical engineering calculations, for they relate the partial pressure of a gas to the total pressure of the mixture. Since the fractions n_1/n_t, n_2/n_t, and n_3/n_t represent the moles of a particular constituent present in the mixture divided by the total number of moles of all gases present, these quantities are called *mol fractions* [1] and are designated by the respective symbols N_1, N_2, N_3, etc. Of necessity the sum of all the mol fractions for a system will have to be unity, namely,

$$N_1 + N_2 + N_3 + \cdots = 1 \tag{14}$$

In terms of these definitions *the partial pressure of any component in a gas mixture is equal to the mol fraction of that component multiplied by the total pressure.* This is true only when the ideal gas law applies to each constituent of the gas mixture.

---

[1] Also referred to occasionally as *pressure* or *volume fractions.*

Amagat's Law of Partial Volumes. A law similar to Dalton's is *Amagat's law of partial volumes.* This law states that *in any gas mixture the total volume may be considered to be the sum of the partial volumes of the constituents of the mixture,* i.e.,

$$V = V_1 + V_2 + V_3 + \cdots \tag{15}$$

where V is the total volume while V_1, V_2, etc., are the partial volumes. By the partial volume of a constituent is meant the volume which that constituent would occupy if present alone at the given temperature and at the *total pressure* of the mixture. By an argument similar to the one employed for partial pressures it is readily shown that, if the ideal gas laws are again applicable, then,

$$V_1 = N_1 V, \qquad V_2 = N_2 V, \text{ etc.} \tag{16}$$

where V_1, V_2, etc., are the partial volumes, N_1, N_2, etc., the mol fractions, and V the total volume at any pressure and temperature.

Dalton's and Amagat's laws are equivalent and hold equally well with gases that approximate ideal behavior, i.e., with gases that are not too close to their condensation temperatures or at too elevated pressures. At high pressures and near their condensation temperatures gases begin to exhibit considerable intermolecular attractions and effects which are no longer general but are specific to the composition and nature of the substances. Under such conditions deviations appear not only from equations (13) and (16), but also from equations (10) and (15). In general the law of partial volumes holds somewhat better than the law of partial pressures at high pressures and low temperatures.

Graham's Law of Diffusion. It has long been known that different gases can diffuse through a tube or escape from a container having a fine opening at different rates dependent on the densities or molecular weights of the gases. The law governing such diffusions was first enunciated by Graham in 1829 and bears his name. This law states that *at constant temperature and pressure the rates of diffusion of various gases vary inversely as the square roots of their densities or molecular weights.* Thus, if we let v_1 and v_2 be the rates of diffusion of two gases, and d_1 and d_2 be their respective densities, then

$$\frac{v_1}{v_2} = \frac{\sqrt{d_2}}{\sqrt{d_1}} \tag{17}$$

Again, since at the same pressure and temperature both gases must have the same molar volume, we have also that

$$\frac{v_1}{v_2} = \frac{\sqrt{d_2 V_m}}{\sqrt{d_1 V_m}} = \frac{\sqrt{M_2}}{\sqrt{M_1}} \tag{18}$$

where M_1 and M_2 are the molecular weights of the two gases.

The Kinetic Theory of Gases. All the principles of gas behavior which have been discussed so far have been arrived at by experiment. The *kinetic theory of gases*, on the other hand, attempts to elucidate the behavior of gases by theoretical means in terms of a postulated "picture" of a gas and certain assumptions regarding its behavior. The theory was first proposed by Bernoulli in 1738, and was considerably elaborated and extended by Clausius, Maxwell, Boltzmann, van der Waals, and Jeans.

The kinetic theory of gases is based on the following fundamental postulates:

(1) Gases are considered to be composed of minute discrete particles called *molecules*. For any one gas all molecules are thought to be of the same mass and size, but to differ in these from gas to gas.

(2) The molecules within a container are not stationary but are believed to be in ceaseless chaotic motion during which they collide with each other and with the walls of the container.

(3) The bombardment of the container walls by the molecules gives rise to the phenomenon we call *pressure*, i.e., the force exerted on the walls per unit area is the average force per unit area which the molecules exert in their collisions with the walls.

(4) Inasmuch as the pressure of a gas within a container does not vary with time at any given pressure and temperature, the molecular collisions must involve no energy loss due to friction. In other words, all molecular collisions are elastic.

(5) At relatively low pressures the average distances between molecules are large compared to molecular diameters, and hence the attractive forces between molecules, which depend on the distance of molecular separation, may be considered negligible.

(6) Again, since the molecules are small compared to the distance between them, they may be considered to a first approximation to be point masses.

(7) Finally, the quantity we call the absolute temperature is a measure of the *average* kinetic energy of all the molecules in a system.

A mathematical analysis of this concept of a gas leads to fundamental conclusions that are directly verifiable by experiment. Consider a cubical container filled with n' molecules of gas, all the same, and all with molecular mass m and velocity u. This velocity u may be resolved into its three components along the x, y, and z axes, as is shown in Fig. 3. If we call these velocity components u_x, u_y, u_z, then

$$u^2 = u_x^2 + u_y^2 + u_z^2 \qquad (19)$$

where u is called the *root-mean-square velocity*.[1] Each of these components may now be treated as though a single molecule of mass m were to

---

[1] A root mean square is equal to the square root of the sum of the squares of a series of numbers.

move independently with each of the component velocities in the appropriate directions x, y, or z. The total effect of these independent motions is obtained by combining the velocities according to equation (19).

Suppose now that the molecule of mass m is moving in the x direction to the right with velocity u_x. It will strike the yz plane with a momentum mu_x, and, since the collision is elastic, it will rebound with velocity $-u_x$ and momentum $-mu_x$. Consequently the *change in momentum* per molecule per single collision in the x direction is $mu_x - (-mu_x) = 2\,mu_x$. Before the molecule can strike the

Fig. 3. Resolution of Velocity along x, y, and z Axes

same wall again it must travel to the opposite wall, collide with it, rebound, and return. To do this it must cover the distance $2\,l$, where l is the length of the cube edge. Hence the number of collisions with the right-hand wall which the molecule will experience per second will be $u_x/2\,l$, and thereby the change in momentum per second for the one molecule on the given wall will be

$$(2\,mu_x)\frac{u_x}{2\,l} = \frac{mu_x^2}{l} \tag{20}$$

But the same change in momentum will be experienced also by the same molecule at the other yz plane, so that the total change in momentum per molecule per second in the x direction is twice the quantity in equation (20), or

Change in momentum/second/molecule in x direction $= \dfrac{2\,mu_x^2}{l}$ (21)

A moment's reflection will show that analogous changes in momentum take place in the y and z directions, and that these are given by $2\,mu_y^2/l$ and $2\,mu_z^2/l$ per molecule per second. From these the

Total change in momentum/molecule/second $= \dfrac{2\,mu_x^2}{l} + \dfrac{2\,mu_y^2}{l} + \dfrac{2\,mu_z^2}{l}$

$$= \frac{2\,m}{l}(u_x^2 + u_y^2 + u_z^2)$$

$$= \frac{2\,m}{l}u^2 \tag{22}$$

by equation (19). As there are n' molecules in the cube, the change in momentum per second for all of them will be equation (22) multiplied by n', or

$$\text{Total change in momentum per second} = \frac{2\, n'mu^2}{l} \tag{23}$$

However, the rate of change of momentum is by definition the acting force, f. Again, pressure is the force per unit area. Consequently,

$$P = \frac{f}{A} = \frac{2\, mn'u^2}{lA} \tag{24}$$

where P is the pressure while A is the total area over which the force is applied. For the cube in question $A = 6\, l^2$, and hence,

$$\begin{aligned} P &= \frac{2\, mn'u^2}{6\, l^2 l} \\ &= \frac{mn'u^2}{3\, l^3} \end{aligned} \tag{25}$$

But l^3 is the volume V of the cube, and so

$$P = \frac{mn'u^2}{3\, V}$$

or
$$PV = \frac{1}{3}\, mn'u^2 \tag{26}$$

According to equation (26) the product PV for any gas should equal one-third the mass of all the molecules (mn') multiplied by the square of the root-mean-square velocity. Although this equation was derived on the assumption of a cubical vessel, it can be shown that the same result is obtained no matter what shape of vessel is considered, and consequently the above deduction must be perfectly general. Equation (26) is the fundamental relation of the kinetic theory of gases.

Deductions from Kinetic Theory of Gases

(a) *Boyle's Law.* We have seen that one of the fundamental postulates of the kinetic theory is the direct proportionality between kinetic energy of the molecules, i.e., $1/2\, mn'u^2$, and the absolute temperature, namely, that

$$\frac{1}{2}\, mn'u^2 = kT \tag{27}$$

where k is a proportionality constant. If now equation (26) is multiplied and divided by 2, we have

$$PV = \frac{2}{3}\left(\frac{1}{2}\, mn'u^2\right)$$

and hence, on insertion of equation (27),

$$PV = \frac{2}{3}\left(\frac{1}{2}\,mn'u^2\right)$$
$$= \frac{2}{3}\,kT \tag{28}$$

When the temperature is constant, equation (28) becomes thus $PV = 2/3\ kT = $ constant, or Boyle's law.

(b) **Charles's Law.** This law holds at constant pressure. If this condition is imposed on equation (28), we get

$$V = \left(\frac{2\,k}{3\,P}\right)T$$
$$= k'T \tag{29}$$

which is a statement of Charles's law.

(c) **Avogadro's Principle.** In 1811 Avogadro enunciated the principle that *equal volumes of all gases at the same pressure and temperature contain equal numbers of molecules.* This principle is readily deducible from the kinetic theory of gases. Since the volumes and pressures are equal, $P_1V_1 = P_2V_2$ for two different gases, and hence it follows from equation (26) that

$$\frac{1}{3}\,n_1'm_1u_1^2 = \frac{1}{3}\,n_2'm_2u_2^2$$

Again, as the temperature is also constant, the average kinetic energy per molecule must be the same or

$$\frac{1}{2}\,m_1u_1^2 = \frac{1}{2}\,m_2u_2^2$$

Inserting the latter relation into the preceding, we see that

$$n_1' = n_2' \tag{30}$$

which is a statement of Avogadro's principle.

The actual number of molecules in a gram-mole of any gas is an important physical constant known as *Avogadro's number,* symbol N. This constant may be arrived at by a number of methods. The best present value for this quantity is 6.023×10^{23} molecules per gram-mole. Once this constant is available the mass of any particular molecule can readily be computed by merely dividing the molecular weight of the substance by Avogadro's number. Thus, since the molecular weight of oxygen is 32, the mass of an individual molecule must be

$$M_{O_2} = \frac{32}{6.023 \times 10^{23}} = 5.31 \times 10^{-23}\ \text{g/molecule}$$

(d) *Graham's Law of Diffusion.* Like the other gas laws, Graham's law follows readily from the kinetic theory of gases. Since at constant volume and pressure for two different gases

$$\frac{1}{3} n_1' m_1 u_1^2 = \frac{1}{3} n_2' m_2 u_2^2$$

then,

$$\frac{u_1^2}{u_2^2} = \frac{m_2 n_2'}{m_1 n_1'}$$

and

$$\frac{u_1}{u_2} = \sqrt{\frac{m_2 n_2'}{m_1 n_1'}} \tag{31}$$

Further, if $n_2' = n_1' = N$, then

$$\frac{u_1}{u_2} = \sqrt{\frac{m_2 N}{m_1 N}} = \sqrt{\frac{M_2}{M_1}} \tag{32}$$

Again, since at constant temperature and pressure the molar volumes are identical, we have also

$$\frac{u_1}{u_2} = \sqrt{\frac{\dfrac{M_2}{V}}{\dfrac{M_1}{V}}} = \sqrt{\frac{d_2}{d_1}} \tag{33}$$

where $d_2 = M_2/V$ and $d_1 = M_1/V$ are the densities of the two gases. Equations (32) and (33) are identical with (17) and (18), and are, of course, statements of Graham's law.

All these deductions point to the fact that the theoretical relation $PV = 1/3 \, n'mu^2$ is in agreement with the empirical ideal gas law $PV = nRT$. Consequently we may write without further hesitation that

$$PV = \frac{1}{3} n'mu^2 = nRT$$

and, since $n' = nN$,

$$PV = \frac{1}{3} n(Nm)u^2 = nRT$$

$$= \frac{nMu^2}{3} = nRT \tag{34}$$

where $M = Nm$ is the molecular weight of the gas in question, and n is the number of *moles* of gas in the volume V at pressure P and temperature T.

Further Deductions from the Kinetic Theory. The value of any theory lies not only in its ability to account for known experimental facts but also in its suggestiveness of new modes of attack. In this respect

the kinetic theory of gases has been very fruitful. We have seen that equation (26), a direct consequence and expression of the theory, gives all the laws of ideal gas behavior. At the same time, however, many other highly important relations can be deduced from it, some of which are outlined below.

(a) *The Velocity of Gas Molecules.* According to the kinetic theory all molecules at the same temperature must have the same average kinetic energy, i.e.,

$$\frac{1}{2} m_1 u_1^2 = \frac{1}{2} m_2 u_2^2 = \frac{1}{2} m_3 u_3^2, \text{ etc.}$$

It follows, therefore, that the higher the mass of a molecule the more slowly must it be moving. It is of considerable interest to ascertain the actual velocity with which various molecules move. From equation (34) we have that

$$\frac{1}{3} n M u^2 = n R T$$

and hence,
$$u = \sqrt{\frac{3\,RT}{M}} \tag{35a}$$

Again, since $RT = PV/n$, and $nM/V = d$, the density of the gas in question at temperature T and pressure P, equation (35a) may be written also as

$$u = \sqrt{\frac{3\,P}{d}} \tag{35b}$$

By either of these equations the root-mean-square velocity of a gas may be calculated from directly measurable quantities. In doing this R must be expressed in ergs per degree per mole, P in dynes per square centimeter, and the density in grams per cubic centimeter. With these units u will be given in centimeters per second.

To calculate the velocity of hydrogen molecules at 0° C we know that $R = 8.315 \times 10^7$ ergs per mole per degree, $T = 273.15$, and $M = 2.016$. Hence equation (35a) yields for u

$$u = \sqrt{\frac{3\,RT}{M}}$$
$$= \sqrt{\frac{3 \times 8.315 \times 10^7 \times 273.15}{2.016}}$$
$$= 184{,}000 \text{ cm/second}$$
$$= 68 \text{ miles/minute}$$

Since hydrogen is the lightest of all elements, this tremendously high velocity represents an upper limit for rates of molecular motion. For all other molecules the speeds will be lower in accordance with Graham's law. Thus for sulfur dioxide, with $M = 64$, the velocity at $0°$ C would be

$$\frac{u_{SO_2}}{u_{H_2}} = \sqrt{\frac{M_{H_2}}{M_{SO_2}}}$$

$$\frac{u_{SO_2}}{68} = \sqrt{\frac{2}{64}}$$

$$u_{SO_2} = 12 \text{ miles/minute}$$

(**b**) *The Kinetic Energy of Translation.* The only type of energy we have ascribed thus far to gas molecules is that due to molecular motion along three coordinate axes, i.e., *kinetic energy of translation.* The amount of this energy is again deducible from equation (34). Since from this equation

$$\frac{1}{3} nM u^2 = nRT$$

and since the kinetic energy, E_k, is given by

$$E_k = \frac{1}{2} nM u^2$$

then $$E_k = \frac{3}{2} \left(\frac{1}{3} nM u^2 \right)$$

$$= \frac{3}{2} nRT \tag{36a}$$

for n moles, or $$E_k = \frac{3}{2} RT \tag{36b}$$

per mole. Consequently the translational energy of an ideal gas is completely independent of the nature or pressure of the gas, and depends only on the absolute temperature. At, say, $300°$ K all ideal gases will thus contain per mole

$$E_k = \frac{3}{2} R(300)$$

$$= 450 R$$

$$= 900 \text{ cal (approx.)}$$

of translational kinetic energy.

(**c**) *Distribution of Molecular Velocities.* For convenience of treatment all molecules in a given gas were considered to be composed of molecules moving with a constant root-mean-square velocity u. Actually, however, all molecules do not possess a constant velocity, for as a

result of collisions a redistribution of both energy and velocity takes place. Maxwell, utilizing probability considerations, has in fact shown that the actual distribution of molecular velocities in a gas depends both on the temperature and molecular weight of a gas, and follows in general the course shown in Fig. 4.
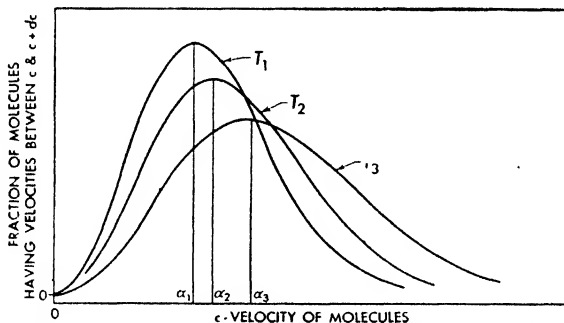

Fig. 4. Distribution of Molecular Velocities in a Gas

In this figure the ordinate represents schematically the fraction of molecules out of a total aggregation which will have a velocity between c and $c + dc$, while the abscissa has the velocity c. The various plots indicate different temperatures which increase in the order T_1, T_2, T_3. From these plots it may be seen that the probability of a molecule being absolutely motionless at any instant is very small. Further, for incidence of velocities greater than zero the probability increases with c, passes through a maximum, and then falls away more or less rapidly toward zero again for very high rates of motion. It is evident, therefore, that both very low and very high speeds are highly improbable, and that most of the molecules in a gas have velocities grouped quite closely about the *most probable velocity* corresponding to the peak of the curve at each temperature. The most probable velocity is in any gas not a constant, but shifts toward higher values of c with increase in temperature; i.e., at higher temperatures higher velocities are more probable than at low.

Mathematical analysis shows that the most probable velocity, α, is not equal either to the root-mean-square velocity u or the average velocity of all the molecules v. If we designate by c_1, c_2, $c_3 \cdots c_n$ the individual velocities of n molecules in a gas, then the average velocity v is defined as

$$v = \frac{c_1 + c_2 + c_3 + \cdots c_n}{n}$$

while the root-mean-square velocity is

$$u = \sqrt{\frac{c_1^2 + c_2^2 + c_3^2 + \cdots c_n^2}{n}}.$$

Kinetic theory arguments reveal that these various velocities are related by the equations

$$v = 0.921\ u$$

$$\alpha = \sqrt{\frac{2}{3}}\,u$$

and hence, on substitution of the value of u from equation (35a), we have

$$u = \sqrt{\frac{3\ RT}{M}} \tag{35a}$$

$$v = 0.921\ \sqrt{\frac{3\ RT}{M}} \tag{37}$$

$$\alpha = \sqrt{\frac{2\ RT}{M}} \tag{38}$$

or $\qquad \alpha : v : u = 1 : 1.128 : 1.224$

(d) Frequency of Collisions and Mean Free Path. It can be shown by a fairly simple argument that in a gas containing n identical molecules per cubic centimeter, the number of molecules with which a *single* gas molecule will collide per second is

$$\sqrt{2}\ \pi v\sigma^2 n$$

where v is the average molecular velocity in centimeters per second and σ the molecular diameter in centimeters. Hence the total number of colliding molecules per cubic centimeter per second, Z, must be n times this quantity, or

$$\begin{aligned} Z &= n(\sqrt{2}\ \pi v\sigma^2 n) \\ &= \sqrt{2}\ \pi v\sigma^2 n^2 \end{aligned} \tag{39}$$

Further, since each collision involves two molecules, the number of molecular collisions occurring in each cubic centimeter per second, N_c, will be one-half this number, namely,

$$N_c = \frac{Z}{2} = \frac{1}{\sqrt{2}}\ \pi v\sigma^2 n^2 \tag{40}$$

Another important quantity in kinetic theory considerations is the average distance a molecule traverses before colliding, or the *mean free path, l*, as it is commonly referred to. If a molecule has an average velocity v cm per second, and if within this period it experiences, as we have seen, $\sqrt{2}\ \pi v\sigma^2 n$ collisions, then the average distance between collisions, or mean free path, must be

$$l = \frac{v}{\sqrt{2}\ \pi v\sigma^2 n}$$

$$= \frac{1}{\sqrt{2}\ \pi\sigma^2 n} \tag{41}$$

The quantities N_c, Z, and l are readily calculable as soon as the molecular diameters σ are available. These are usually obtained from gas viscosity measurements, for according to kinetic theory the viscosity coefficient η is given approximately by the relation

$$\eta = \frac{1}{3} vld \qquad (42)$$

where d is the density and the other symbols have the usual significance. Knowing the mean free path l, σ may be calculated by equation (41). The question of gas viscosity will be discussed more fully toward the end of this chapter.

Applicability of the Ideal Gas Laws. The concordance between the empirical generalizations embodied in the expression $PV = nRT$ and the deductions of the kinetic theory of gases lends considerable credence to our conception of the nature of gases and their behavior. However, there still remains the question of how completely and accurately can the expression $PV = nRT$ reproduce the actual P-V-T relations of all gases. To test this point we may resort to the fact that at constant temperature the combined gas law reduces to $PV = nRT = $ constant. Hence, as long as T does not vary, the product PV for a given quantity of gas should remain the same at all pressures. A plot of PV vs. P at constant T should yield, therefore, a straight line parallel to the abscissa.
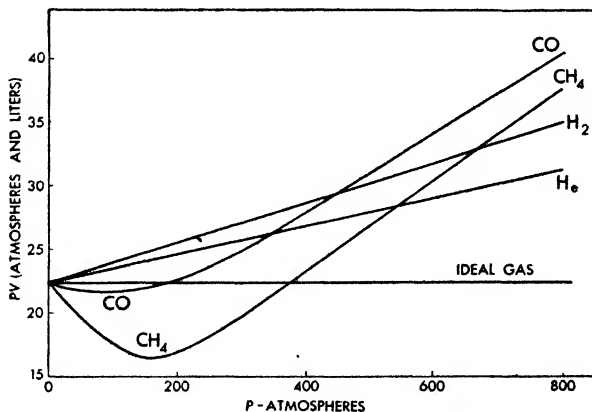

Fig. 5. PV vs. P Plot for Several Gases at $0°$ C ($n = 1$)

Such a plot of PV vs. P constructed from actual data for several typical gases at $0°$ C is shown in Fig. 5. The fact immediately apparent is that the expression $PV = k$ is not obeyed over most of the pressure range shown. The curves obtained are in general of two types. One, including only hydrogen and helium here, starts at the value of PV

demanded by $PV = nRT$ for the temperature in question and increases continually with pressure. In every case the product PV is greater than demanded by theory, and hence these gases are less compressible than expected, or, in other words, they are superperfect. On the other hand, in the second type the plot starts again at the same point as before, but now the product PV decreases at first with pressure, passes through a minimum characteristic of each gas and the temperature, and then increases to values which may rise appreciably above the theoretical. In the range of pressures where PV is less than nRT the gases are more compressible than ideal gases, and hence here the gases are underideal or underperfect.

The type of curve which a particular gas will exhibit at a given temperature depends on how far it is removed from its critical temperature.[1]
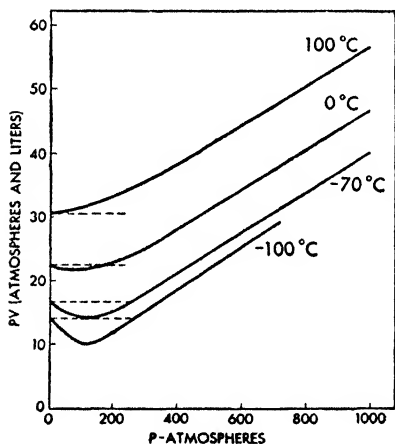

The closer a given temperature is to the critical temperature the lower will be the minimum in the curve and the greater will be the deviations from the ideal gas law. And conversely, the farther away a gas is from its critical temperature the shallower will be the minimum, and the more closely will the gas approximate ideal behavior at the lower pressures. In fact, at considerable distances from the critical, the minimum disappears, and a gas passes over to the behavior exhibited by hydrogen and helium at 0° C and above. This may be seen from the data for nitrogen shown in Fig. 6.

Fig. 6. Compressibilities of Nitrogen at Several Temperatures $(n = 1)$

From this description it is to be anticipated that at sufficiently low temperatures even hydrogen and helium will exhibit minima in their PV vs. P plots, and this is actually the case.

The highly individualistic behavior exhibited by various gases indicates that in order to represent their $P–V–T$ relations equations of state, i.e., equations involving P, V, and T, would be required which would contain not only these variables, but also terms making allowance for the specific forces operative in each gas. However, $P–V–T$ studies on gases at low pressures do show that when the pressures are lowered gases begin to approximate more closely the ideal gas law, and, furthermore, the lower the pressure the better is the agreement between the observed PV product and that calculated from the combined gas law. At these

[1] The highest temperature at which a gas may be liquefied.

low pressures all gases lose their individualistic behavior and merge to obey the simple and general expression obtained from the kinetic theory of gases. For this reason the expression $PV = nRT$ is considered to be a *limiting law* only, a law which gases obey strictly only when they are diluted highly enough so that the volume of the molecules themselves is negligible compared to the total volume, and the intermolecular attractive forces are too feeble to exercise any influence on the pressure of the gas. It may be concluded, therefore, that a gas becomes more ideal as the pressure is lowered, and will become completely ideal as the pressure approaches zero.

How far this concordance between the ideal gas law and observation will extend into the range of higher pressures depends on the nature of the gas and the temperature. For gases which are permanent at ordinary temperatures, i.e., which are above their critical temperatures, such as hydrogen, nitrogen, oxygen, and helium, this concordance may extend within 5 per cent or so up to pressures as high as 50 atmospheres. On the other hand, with easily condensible gases, such as carbon dioxide, sulfur dioxide, chlorine, and methyl chloride, discrepancies as large as 2 or 3 per cent may appear at 1 atmosphere pressure. The use of the ideal gas law for such gases is considerably limited, therefore, when fairly precise calculations are required. In any case, before using the ideal gas law at any appreciable pressure it is always advisable to consider the nature of the gas in question and the distance it is removed from its critical temperature. The greater this distance, the wider in general will be the pressure range over which calculations can be made within a given accuracy.

The van der Waals Equation of State. Because of the deviation of real gases from the behavior demanded by the ideal gas law, many attempts have been made to set up equations of state which will reproduce more satisfactorily the P–V–T relations of gases. Of these equations one of the earliest and best known is that of van der Waals.

The van der Waals equation differs from the ideal gas law in that it makes allowance both for the volume occupied by the molecules themselves and for the attractive forces between them. In the simple kinetic theory, upon which the ideal gas laws are predicated, the molecules are considered to be point masses occupying zero volume. However, although such an assumption is not serious when the molecules are far apart, i.e., at low pressures, at high pressures, when the molecules are close together, the volume of the molecules themselves may become a sizable fraction of the total volume. Under such conditions the "free space" available for compression becomes considerably less than the total volume, and the latter cannot be taken as the space through which a gas may be compressed.

To make this correction necessitated by the finite dimensions of the molecules, let b be the effective volume of the molecules in *one* mole of gas and V the volume of n moles of gas. In this total volume that occupied by the molecules themselves will be thus nb, and hence the volume available for compression will be not V but $(V - nb)$. Since the latter is the "free space," it should be substituted for V in the ideal gas law. It may be anticipated that b will be characteristic and different for each gas.

The second factor of which van der Waals took cognizance is the attractive force operative between molecules. These attractive forces are inverse functions of some fairly high power of the distance between molecules. When a gas is compressed, the molecules are brought closer together and the force of attraction increases; conversely, as the molecules are separated by decreasing the pressure, the forces fall off rapidly. The effect of these attractive forces is to produce within a container a pressure which is *smaller* than that which would be produced if the molecules did not experience attractions toward each other. The manner in which this decrease in pressure occurs can be visualized from the following argument.

Consider a wall of the container which is being bombarded by gaseous molecules. The force per unit area due to this bombardment is the observed pressure. When the gas molecules are not constrained by attractions for each other, they will bombard the walls with the full force due to their outward motion. If, however, under the same conditions a molecule moving outward is subjected by molecular attraction to an inward "pull," some of the energy of its outward motion will have to be employed to overcome the inward attraction. Such a molecule will not strike the wall with as high a force as if it were not "dragged back" by the other molecules within the gas, and consequently the pressure resulting from the bombardment will be lessened by an amount P'. The observed pressure, P, will thus be less than the ideal pressure, P_i, by the amount P', or

$$P = P_i - P'$$

Since in the expression $P_i V = nRT$ the pressure P_i refers to the ideal pressure, we must substitute for it its value from the expression given above, or $P_i = (P + P')$. If we combine this corrected pressure with the expression for the corrected volume we obtain, instead of $PV = nRT$, the equation

$$(P + P')(V - nb) = nRT \qquad (43)$$

van der Waals indicated that the magnitude of the pressure correction P' for n moles of gas present in volume V is given by

$$P' = \frac{n^2 a}{V^2}$$

where a is a constant characteristic of each gas and independent of pressure and temperature. It is for each gas a measure of the magnitude of the intermolecular attractive forces within the gas. If this expression for P' is substituted in equation (43), we get

$$\left(P + \frac{n^2 a}{V^2}\right)\left(V - nb\right) = nRT \tag{44}$$

This is the celebrated equation of state which was first developed by van der Waals in 1873 and which bears his name.

Use of van der Waals' Equation. In applying van der Waals' equation care must be exercised in the choice of appropriate units, particularly for the constants a and b. Since $n^2 a / V^2$ must represent a pressure, the units of a must be pressure $\times$ (volume)^2/(moles)^2, i.e., atmospheres $\times$ (liters)^2/(moles)^2, or atmospheres $\times$ (cc)^2/(moles)^2. In any event, the units used must be the same as those of P and V, and this applies also to R. In turn, b is a volume and must correspond to the units of V. T, of course, is in ° K.

The use of the equation can best be illustrated with an example. Suppose it is desired to calculate by van der Waals' equation the pressure at which 2 moles of ammonia will occupy a volume of 5 liters at 27° C. For ammonia, $a = 4.17$ atm $\times$ liter^2, while $b = 0.0371$ liter per mole. Hence,

$$\begin{aligned} P &= \frac{nRT}{V - nb} - \frac{n^2 a}{V^2} \\ &= \frac{2(0.0821)300.2}{5 - 2(0.0371)} - \frac{(2)^2 \times 4.17}{(5)^2} \\ &= 9.33 \text{ atm} \end{aligned}$$

The corresponding pressure calculated from the ideal gas law is 9.86 atmospheres.

Table 2 lists the van der Waals constants a and b for a number of gases. Such gases as carbon disulfide, ammonia, sulfur dioxide, chloroform, etc., which are easily condensible, have relatively high values of a, indicating strong intermolecular attractions. On the other hand, for the permanent gases such as argon, carbon monoxide, helium, and hydrogen, the a values are considerably lower, and hence in these the intermolecular forces are considerably weaker.

The van der Waals equation is much more accurate than the simple ideal gas law and is valid over a much wider pressure range, as may be

TABLE 2

VAN DER WAALS CONSTANTS FOR VARIOUS GASES

$$\left[a \text{ in } \frac{\text{(atm)(liters)}^2}{\text{(mole)}^2}; b \text{ in } \frac{\text{(liters)}}{\text{(mole)}} \right]$$

Gas	Formula	a	b
Acetylene	C_2H_2	4.39	0.0514
Ammonia	NH_3	4.17	0.0371
Argon	A	1.35	0.0322
Carbon dioxide	CO_2	3.59	0.0427
Carbon disulfide	CS_2	11.62	0.0769
Carbon monoxide	CO	1.49	0.0399
Carbon tetrachloride	CCl_4	20.39	0.1383
Chlorine	Cl_2	6.49	0.0562
Chloroform	$CHCl_3$	15.17	0.1022
Ethane	C_2H_6	5.49	0.0638
Ethyl ether	$(C_2H_5)_2O$	17.38	0.1344
Ethylene	$CH_2{=}CH_2$	4.47	0.0571
Helium	He	0.034	0.0237
Hydrogen	H_2	0.244	0.0266
Hydrogen bromide	HBr	4.45	0.0443
Hydrogen chloride	HCl	3.67	0.0408
Methane	CH_4	2.25	0.0428
Neon	Ne	0.211	0.0171
Nitric oxide	NO	1.34	0.0279
Nitrogen	N_2	1.39	0.0391
Nitrogen dioxide	NO_2	5.28	0.0442
Nitrous oxide	N_2O	3.78	0.0442
Oxygen	O_2	1.36	0.0318
Sulfur dioxide	SO_2	6.71	0.0564
Water	H_2O	5.46	0.0305

seen from Table 3. However, under extreme conditions, such as temperatures near the critical and at very high pressures, its predictions deviate considerably in many instances from experimentally observed values. It is very doubtful whether it is justifiable to consider a and b as constants independent of pressure and temperature. In fact, in order to fit the equation to experimental data with a relatively high order of fidelity it is necessary to choose different values of a and b over different ranges of pressure and temperature.

Other Equations of State. A large number of equations of state other than the van der Waals have been proposed to represent the P–V–T relations of gases. Some of these are based to some extent on theoretical considerations, while others are entirely empirical. We shall consider now several of the more important of these equations.

TABLE 3

COMPARISON OF IDEAL GAS LAW AND VAN DER WAALS' EQUATION AT 100° C

Observed P (atm)	Hydrogen				Carbon Dioxide			
	P calc. Ideal	% Deviation	P calc. van der Waals	% Deviation	P calc. Ideal	% Deviation	P calc. van der Waals	% Deviation
50	48.7	− 2.6	50.2	+ 0.4	57.0	+ 14.0	49.5	− 1.0
75	72.3	− 3.6	75.7	+ 0.9	92.3	+ 17.3	73.3	− 2.3
100	95.0	− 5.0	100.8	+ 0.8	133.5	+ 33.5	95.8	− 4.2

The Kammerlingh Onnes Equation of State. This empirical equation expresses PV as a power series of the pressure at any given temperature, namely,

$$PV = A + BP + CP^2 + DP^3 + \cdots \tag{45}$$

P is the pressure, generally in atmospheres, and V is the *molar* volume in liters or cubic centimeters. The coefficients A, B, C, etc., are known respectively as the first, second, third, etc., *virial coefficients*. At very low pressures only the first of these coefficients is significant, and it is equal essentially to RT. At higher pressures, however, the others as well are important and must be considered. In general the order of significance of the coefficients is their order in the equation. These coefficients, although constant at any given temperature, change in value as the temperature is changed. Of necessity the first virial coefficient A is always positive and increases with temperature. The second coefficient, on the other hand, is negative at low temperatures, passes through zero, and becomes increasingly positive as the temperature is raised. The temperature at which $B = 0$ is known as the *Boyle temperature*, for at this temperature Boyle's law is valid over a fairly wide pressure range.

By using a sufficient number of terms this equation can be fitted to experimental data with a high order of accuracy. The virial coefficients for several gases are shown in Table 4. With these it is possible to calculate PV up to 1000 atmospheres.

The Berthelot Equation. The high-pressure form of this equation is rather difficult to handle. For low pressures the equation reduces to

$$PV = nRT\left[1 + \frac{9\,PT_c}{128\,P_cT}\left(1 - \frac{6\,T_c^2}{T^2}\right)\right] \tag{46}$$

where P, V, R, T, and n have the same meaning as in the ideal gas law, while P_c and T_c are the critical pressure and critical temperature respec-

TABLE 4

VIRIAL COEFFICIENTS OF SOME GASES

(For P in atm, V in liters)

$t°$ C	A	$B \times 10^2$	$C \times 10^5$	$D \times 10^8$	$E \times 10^{11}$
Nitrogen					
− 50	18.312	− 2.8790	14.980	− 14.470	4.657
0	22.414	− 1.0512	8.626	− 6.910	1.704
100	30.619	0.6662	4.411	− 3.534	0.9687
200	38.824	1.4763	2.775	− 2.379	0.7600
Carbon Monoxide					
− 50	18.312	− 3.6878	17.900	− 17.911	6.225
0	22.414	− 1.4825	9.823	− 7.721	1.947
100	30.619	0.4036	4.874	− 3.618	0.9235
200	38.824	1.3163	3.052	− 2.449	0.7266
Hydrogen					
− 50	18.312	1.2027	1.164	− 1.741	1.022
0	22.414	1.3638	0.7851	− 1.206	0.7354
500	63.447	1.7974	0.1003	− 0.1619	0.1050

tively. The meaning of the latter quantities will be discussed more fully in Chapter III. For pressures of about an atmosphere and below this equation is very accurate, and it is consequently very useful in calculating the molecular weights of gases from their densities. Its use will be illustrated in that connection.

The Beattie-Bridgeman Equation of State. This semitheoretical equation of state involving five constants is one of the most accurate equations of state available at present. It may be stated in two forms, one explicit in pressure, the other in molar volume V_m, namely,

$$P = \frac{RT}{V_m} + \frac{\beta}{V_m^2} + \frac{\gamma}{V_m^3} + \frac{\delta}{V_m^4} \tag{47}$$

$$V_m = \frac{RT}{P} + \frac{\beta}{RT} + \frac{\gamma P}{(RT)^2} + \frac{\delta P^2}{(RT)^3} \tag{48}$$

where

$$\beta = RTB_0 - A_0 - \frac{Rc}{T^2} \tag{49a}$$

$$\gamma = - RTB_0 b + A_0 a - \frac{RcB_0}{T^2} \tag{49b}$$

$$\delta = \frac{RB_0 bc}{T^2} \tag{49c}$$

In these relations T is again the absolute temperature and R is the gas constant, while A_0, B_0, a, b, and c are constants characteristic of each gas.

perature and pressure with the gas whose molecular weight is to be determined, and weighed again. The difference in weights represents the weight of gas W in the flask. The volume of the flask is determined by filling it with water or mercury, whose densities are known, and again weighing. From the data thus obtained the molecular weight may be calculated by equation (53).

For more precise work a larger bulb is used to increase the mass of gas and a similar bulb is employed as a counterpoise. The observed weights are also reduced to vacuo.

Dumas' Method for Determination of Vapor Densities. This method is used to determine the molecular weights in the vapor phase of readily volatile liquids. A retort-shaped bulb, having a small opening drawn to a capillary, is first weighed full of air. A sample of several cubic centimeters of the liquid in question is drawn into the bulb by cooling it with the tip below the surface of the liquid, and the bulb is then immersed in a bath whose temperature is above the boiling point of the liquid. The boiling is permitted to proceed until the vapors of boiling liquid have expelled all the air from the bulb, and the liquid in the flask has completely vaporized. The flask is then sealed, cooled to room temperature, and weighed. The volume of the bulb is determined as in Regnault's method. The pressure of the vapor when the bulb is sealed is the same as atmospheric, while the temperature is that of the bath. The weight of vapor, after corrections for buoyancy, is obtained from the following equation:

$$W_{vapor} = W_{(bulb + vapor)} - W_{(bulb + air)} + W_{air} \qquad (55)$$

W_{air} is obtained by multiplying the volume of the flask by the density of the air. Knowing P, V, T, and W_{vapor}, the molecular weight of the liquid in the vapor phase may be calculated as before, or the density of the vapor may be calculated from $d = W_{vapor}/V$, and equation (54) can be used to find M.

The Victor Meyer Method for Vapor Densities. This method, suggested by Victor Meyer, serves the same purpose as the Dumas method for the determination of vapor densities, but is considerably simpler and more flexible. A sketch of the apparatus is shown in Fig. 8. It consists of an inner tube B, approximately 50 cm long, which is surrounded by a jacket A, partly filled as indicated with a liquid whose boiling point is at least 30° higher than that of the substance to be studied. The function of the outer jacket is to keep the temperature of the inner tube constant by boiling the liquid in A throughout a run. Inside the inner tube, in turn, is another tube C, open at the bottom, down which passes a metal or glass rod, anchored with rubber tubing at the top in the manner shown and fitted with a hook at the bottom.

The outlet from B communicates with a gas burette G, filled either with water, in which case correction for the aqueous pressure must be applied, or preferably mercury. L is a leveling bulb to permit adjustment of gas pressure in G to that of the atmosphere.
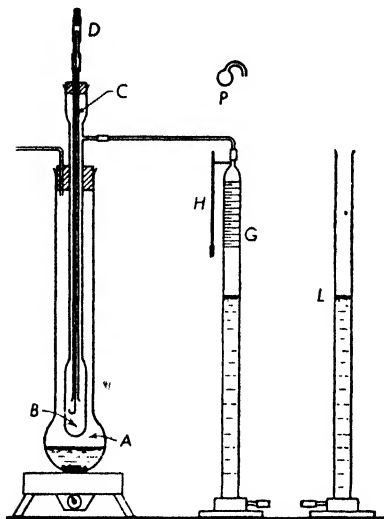
The liquid whose molecular weight is to be determined is enclosed in a small glass ampoule with finely drawn tip, P. This ampoule is first weighed empty, next enough of the liquid is drawn in to yield 40 to 60 cc of vapor, and the bulb is sealed carefully in a flame and weighed again. The difference between the first and second weighings gives the weight of the liquid W to be vaporized. This ampoule is hung then on the hook projecting from C, and the entire apparatus assembled as shown in the figure.

Fig. 8. Victor Meyer Apparatus

To make a measurement, the liquid in A is brought to boiling and kept there for the entire run. When thermal equilibrium has been established, the levels in G and L are equalized and the burette reading is taken. Next the ampoule is smashed by pulling upward on the rod at D so as to bring the neck of the ampoule up against the bottom of C. With the bulb broken the liquid vaporizes, and the vapors generated displace air from the bottom of B into the gas burette G. The volume of air thus displaced is equal to the volume of the vapors formed at the temperature of the inner tube. Once in the gas burette the air cools to room temperature, and its volume can be measured by again reading the burette. Provided the levels in G and L are equalized, the pressure of this air is the same as that of the atmosphere outside the burette, while the temperature is that read on the thermometer H. The volume of displaced air thus obtained, i.e., final minus initial burette readings, is equal to the volume which the vapors of the liquid would occupy if they could be cooled to the temperature of the room and atmospheric pressure. Having measured in this manner the weight of liquid W, and its volume as a *vapor* at room temperature T and barometric pressure P, the density of the vapor and its molecular weight may readily be calculated from the observed data.

To illustrate the method of calculation, consider the following example. In determining the molecular weight of ethyl alcohol the data obtained were:

Weight of liquid taken = 0.1211 g
Volume of air measured over water = 67.30 cc
Temperature = 28.0° C
Atmospheric pressure = 755.2 mm Hg (corrected)
Aqueous tension of water at 28° C (from tables) = 28.3 mm Hg

Since the total pressure in the burette is the sum of the pressures of vapor P_V and that of water P_{H_2O}, and since this total is equal to atmospheric pressure, then

$$P_V = P_{atm} - P_{H_2O}$$
$$= 755.2 - 28.3$$
$$= 726.9 \text{ mm Hg}$$

Inserting this value of P_V along with those of W, T, and V into equation (53), the molecular weight of ethyl alcohol follows as

$$M = \frac{WRT}{PV}$$
$$= \frac{0.1211 \times 82.06 \times 301.2}{\left(\dfrac{726.9}{760}\right) 67.30}$$
$$= 46.5$$

The molecular weight obtained from atomic weights is **46.07**. Likewise the density may be calculated from equation (54).

Exact Determination of Vapor Densities and Molecular Weights. The molecular weights calculated from the ideal gas laws are, even with good data, only approximate. The reason is that already at atmospheric pressure the ideal gas laws fail to represent accurately the behavior of the vapors. Measurements on vapors are usually made at temperatures close to the boiling point, and under such conditions deviations from ideal behavior are to be expected. For most purposes an approximate molecular weight when used in conjunction with an empirical chemical formula will suffice to establish the exact molecular weight of a substance in the gas phase. Occasionally, however, as in atomic weight determinations, an exact molecular weight is desired, and this must be obtained from either a more precise gas equation or by special treatment of the ideal gas laws.

When the constants a and b of a substance are known, use of van der Waals' equation will give better concordance between observed and calculated values of the molecular weight. For the purpose at hand, however, the Berthelot equation is more convenient and gives good results. It can be used, of course, only when the critical temperature and pressure of the substance are available. The Berthelot equation (46) is read-

ily transformable to yield either the molecular weight or the density. Since $n = W/M$, equation (46) gives for M

$$M = \left(\frac{W}{V}\right)\left(\frac{RT}{P}\right)\left[1 + \frac{9\,PT_c}{128\,P_cT}\left(1 - \frac{6\,T_c^2}{T^2}\right)\right] \qquad (56)$$

Further, since $W/V = d$, the density, equation (56) may also be written as

$$M = \frac{dRT}{P}\left[1 + \frac{9\,PT_c}{128\,P_cT}\left(1 - \frac{6\,T_c^2}{T^2}\right)\right] \qquad (57)$$

from which the density follows when M is known or vice versa.

The higher accuracy of the Berthelot equation can be illustrated with the following data on methyl chloride. For methyl chloride, $T_c = 416.2°$ K, $P_c = 65.8$ atmospheres, while the density at standard conditions is 2.3076 grams per liter. Hence, by equation (57),

$$M = \frac{2.3076 \times 0.08206 \times 273.2}{1}\left[1 + \frac{9 \times 1 \times 416.2}{128 \times 65.8 \times 273.2}\left(1 - 6\frac{(416.2)^2}{(273.2)^2}\right)\right]$$
$$= 50.62$$

as against the theoretically calculated 50.48. Using the same data and the ideal gas law, the molecular weight obtained is 51.71.

A means of obtaining exact molecular weights is the method of *limiting densities*. This method, which gives excellent results, is based upon the fact that as zero pressure is approached the ideal gas laws become exact for all gases. The densities of a gas or vapor are determined at a given temperature at atmospheric pressure and at several other pressures below one atmosphere. The ratio d/P or W/PV is then plotted against P. If the vapor or gas were ideal, this ratio would remain constant for all pressures, for

$$P = \frac{W}{VM}RT$$

$$P = \frac{d}{M}RT$$

$$\frac{d}{P} = \frac{M}{RT} = \text{constant} \qquad (58)$$

However, since this is not true for real gases, the ratio d/P changes with decreasing pressure. Fortunately the plot is practically linear and can be extrapolated to zero pressure without any difficulty. At zero pressure the limiting ratio d/P is that for the ideal gas and so

$$\left(\frac{d}{P}\right)_{P\to 0} = \frac{M}{RT}$$

and
$$M = RT\left(\frac{d}{P}\right)_{P\to 0} \qquad (59)$$

This method can be illustrated with the data on hydrogen bromide shown in Table 6, while the plot of d/P vs. P is shown in Fig. 9. The
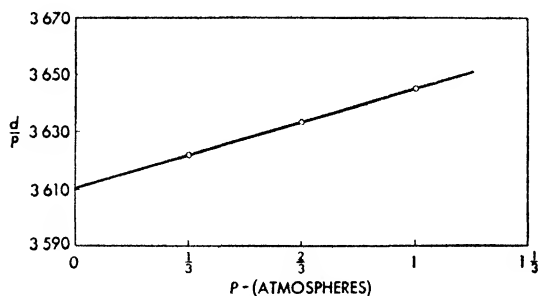

Fig. 9. Plot of $\dfrac{d}{P}$ vs. P for HBr at 0° C

extrapolated value of d/P is 3.6108 grams per liter per atmosphere at 0° C. Hence the molecular weight of hydrogen bromide is

$$M = 3.6108 \times 0.08206 \times 273.15 = 80.94$$

The value calculated from atomic weights is 80.92.

TABLE 6

DENSITIES OF HBr AT VARIOUS PRESSURES (0° C)

P (atm)	d (g/liter)	d/P
1	3.6444	3.6444
2/3	2.4220	3.6330
1/3	1.2074	3.6222
0	—	3.6108 (extp'd)

Results of Vapor Density Measurements. The measurement of the vapor densities of a large number of substances shows that the molecular weight of these substances in the gas phase over a certain temperature interval is what would be expected from their simple formula. Among these may be mentioned ammonia, carbon dioxide, hydrogen, nitrogen, carbon monoxide, methyl chloride, methyl fluoride, ethyl ether, methyl ether, carbon tetrachloride, chloroform, carbon disulfide, acetone. There are other substances, however, which exhibit a highly anomalous behavior. These may be segregated into two groups: (1) those which exhibit vapor densities, and consequently molecular weights, very much higher than would be expected on the basis of their simple formulas, and (2) those which exhibit vapor densities much lower than that expected from their simple formulas. All these abnormalities are very much greater than can be accounted for by either experimental uncertainty or deviation from ideal behavior.

The substances exhibiting abnormally high vapor densities are considered to be associated in the vapor phase, i.e., the molecule is considered to be composed of more than a single structural unit. In line with this view is the fact that the calculated molecular weight is usually a whole-number multiple of the simple formula. Thus aluminum chloride, $AlCl_3$, is shown in the vapor phase to be $(AlCl_3)_2$ or Al_2Cl_6. Similarly ferric chloride is Fe_2Cl_6, beryllium chloride Be_2Cl_4, and gallium chloride Ga_2Cl_6. Sulfur is another substance which shows different stages of association in the gas phase at different temperatures.

The question of the abnormally low vapor densities disturbed chemists in the nineteenth century for many years. Although it is fairly easy to see that a molecule may involve several simple units, and thus give a high molecular weight, it is more difficult to understand how the molecular weight of a substance can be lower than that demanded by the simple empirical formula. For a while the validity of Avogadro's hypothesis was questioned, but if the validity of this principle is granted, the conclusion is that the substances exhibiting abnormally low vapor densities break down or dissociate in the vapor phase under the influence of heat into simpler substances, leading thus to a greater number of particles and a lower density for any given pressure. Seville in 1857 and Kopp in 1864 actually confirmed this for ammonium chloride. They showed that the vapor of ammonium chloride, which exhibits an abnormally low density, contained ammonia and hydrogen chloride, and concluded that on heating the reaction

$$NH_4Cl = NH_3 + HCl$$

took place. This explanation holds as well for all the other substances in this category. Thus, phosphorus pentachloride dissociates in the vapor phase as follows,

$$PCl_5 = PCl_3 + Cl_2$$

while nitrogen tetroxide dissociates according to

$$N_2O_4 = 2\,NO_2$$
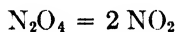
In any instance of abnormally low vapor densities the extent of dissociation is a function of the temperature and pressure. At sufficiently high temperatures these substances may be completely dissociated, while at sufficiently low temperatures they may behave almost normally. In fact, practically all substances can be shown to be abnormal if the temperature is made high enough. Even such a stable compound as carbon dioxide dissociates above $2000°$ C to some extent into carbon monoxide and oxygen. Similarly aluminum chloride at $400°$ C is Al_2Cl_6, at $500°$ a mixture of Al_2Cl_6 and $AlCl_3$, while at $1100°$ it is all $AlCl_3$. If heated

further, $AlCl_3$ will actually dissociate into aluminum and chlorine. Hence when we speak of the molecular weight of a substance in the gas phase, it is very important to keep in mind the temperature to which reference is made.

Specific Heat and Heat Capacity. The specific heat of any substance is defined as the quantity of heat required to raise the temperature of unit weight of the substance 1 degree of temperature. In terms of calories and degrees centigrade, the specific heat of any substance is the number of calories of heat required to raise the temperature of 1 gram of the substance 1 degree centigrade. A calorie of heat, in turn, is defined as the amount of heat required to raise the temperature of 1 gram of water from 15° to 16°C. The temperature is specified because the specific heat of water, like that of most other substances, is not constant but varies with temperature.

Chemical calculations are most frequently made on a molar basis, and for that reason it is very convenient to deal with the specific heat per mole, or *heat capacity* per mole. The heat capacity per mole is the amount of heat required to raise the temperature of 1 mole of a substance 1 degree centigrade. It is equal, of necessity, to the specific heat per gram multiplied by the molecular weight of the substance.

Two types of specific heats are recognized, depending on whether the substance is heated at constant volume or at constant pressure. When a substance is heated at constant volume, all of the energy supplied goes to increase the internal energy of the substance, and we speak then of the *specific heat* or *heat capacity at constant volume* C_v. On the other hand, when a substance is heated at constant pressure, energy must be supplied not only to increase its internal energy, but also to make possible expansion of the substance against the confining atmospheric pressure. *The specific heat* or *heat capacity at constant pressure* C_p must therefore be larger than that at constant volume by the amount of work which must be performed in the expansion accompanying 1 degree rise in temperature. In liquids and solids, where volume changes on heating are small, this difference between C_p and C_v is usually very slight. With gases, however, where the volume changes with temperature are always large, the difference $C_p - C_v$ is always significant and cannot be disregarded.

Specific Heats of Gases and Kinetic Theory. Some important deductions concerning the specific heats of gases can be made from the kinetic theory of gas behavior. According to equation (36b) the kinetic energy of translation of an ideal gas per mole is

$$E_k = \frac{3}{2} RT \quad .$$

If this is the only type of energy the gas possesses, the energy difference of the gas $(E_{k_2} - E_{k_1})$ between two temperatures T_2 and T_1 is

$$\Delta E = E_{k_2} - E_{k_1} = \frac{3}{2} RT_2 - \frac{3}{2} RT_1$$

$$= \frac{3}{2} R(T_2 - T_1) \text{ per mole}$$

When the temperature difference $T_2 - T_1 = 1$, ΔE becomes the energy required to raise the translational energy of 1 mole of gas 1 degree without involving any external work, or, in other words, the heat capacity per mole at constant volume C_v. Hence we may write

$$C_v = \frac{3}{2} R$$

$$= \frac{3}{2} \times 1.987$$

$$= 2.98 \text{ cal/degree/mole of gas} \tag{60}$$

The kinetic theory predicts, therefore, that C_v for any ideal gas *containing only translational energy* should be approximately 3 calories per mole, and, further, that this heat capacity should be constant and independent of temperature.

A similar prediction can be arrived at with respect to the heat capacity at constant pressure C_p. In view of the preceding considerations it follows that

$$C_p = C_v + w \text{ per degree/mole} \tag{61}$$

where w is the work which must be performed against a confining pressure P when 1 mole of an ideal gas is expanded from a volume V_1 at T_1 to a volume V_2 at $T_2 = T_1 + 1$. The value of w can be obtained readily from the relation

$$w = \int_{V_1}^{V_2} P dV \tag{62}$$

which will be discussed in greater detail in the next chapter. If we differentiate now $PV = RT$ per mole at constant pressure, we have

$$P dV = R dT$$

and on substitution of $R dT$ for $P dV$ in equation (62), we see that

$$w = \int_{V_1}^{V_2} P dV = \int_{T_1}^{T_2} R dT$$
$$= R(T_2 - T_1) \text{ per mole}$$

For $T_2 - T_1 = 1$ this reduces $w = R$ per mole, and hence for an ideal gas

$$C_p = C_v + R \text{ cal/degree/mole of gas} \tag{63}$$

This important conclusion that $C_p - C_v = R$ is valid for all ideal gases, and permits the simple conversion of C_p to C_v or vice versa. Inserting the values of C_v from (60) and of R, we see that for any ideal gas *involving only translational energy* C_p should be

$$C_p = \frac{3}{2}R + R$$
$$= \frac{5}{2}R$$
$$= 4.97 \text{ cal/degree/mole} \qquad (64)$$

Consequently, like C_v, C_p should be constant and independent of temperature for all gases. Again, the ratio C_p/C_v, commonly designated by γ, should also be a constant equal to

$$\gamma = \frac{C_p}{C_v}$$
$$= \frac{\dfrac{5}{2}R}{\dfrac{3}{2}R}$$
$$= 1.67 \qquad (65)$$

In Table 7 are listed values of C_p, C_v, $C_p - C_v$, and γ for various gases at 15° C. It will be observed, first of all, that the requirement $C_p - C_v = R = 1.99$ calories per mole is met fairly well by practically all the gases in the table. Second, the predictions of the kinetic theory that $C_p = 4.97$ and $C_v = 2.98$ calories per mole are borne out by the specific heats of a group of gases which includes besides argon and helium also krypton, xenon, and a number of metallic vapors. However, for all the other gases in the table, the prediction is not valid. A careful inspection of the table reveals that the various gases can be divided into distinct classes based upon their values of γ. The first group, comprising the gases that obey the kinetic theory, has the expected $\gamma = 1.67$. The others, in turn, may be grouped as those with γ equal approximately to 1.40, 1.3, and lower. In no case does γ fall below unity, although it may approach it. Further, the decrease in the value of γ is always associated with an increase in the complexity of the molecules involved. Thus argon and helium with $\gamma = 1.67$ are monatomic, i.e., the molecules are composed only of a single atom of the element. Again, the substances with γ equal to about 1.40, such as oxygen, nitrogen, and chlorine, are diatomic, those with γ equal to about 1.30 triatomic, while all others with γ still lower are more complex. Finally, all substances exhibiting γ values lower than 1.67 also have values of C_p and C_v considerably greater than the predicted $C_p = 5/2 R$ and $C_v = 3/2 R$.

The Heat Capacity of Polyatomic Gases. These "anomalously" high heat capacities suggest that the fundamental assumption made, that the only energy involved in a gas is kinetic energy of translation, is not always correct. A monatomic molecule can execute only translational motion along the coordinate axes, and for such a gas the deductions of the kinetic theory should be valid, as is actually the case. A more complex molecule, however, may be subject not only to translational motion

TABLE 7

HEAT CAPACITIES OF GASES AT 15° C [1]

(per mole)

Gas	Formula	C_p	C_v	$C_p - C_v$	γ
Argon	A	5.00	3.01	1.99	1.66
Helium	He	4.99	3.00	1.99	1.66
Carbon monoxide	CO	6.93	4.93	2.00	1.40
Chlorine	Cl_2	8.14	6.00	2.14	1.36
Hydrogen	$\cdot H_2$	6.82	4.83	1.99	1.41
Hydrogen chloride	HCl	7.07	5.01	2.06	1.41
Nitrogen	N_2	6.93	4.94	1.99	1.40
Oxygen	O_2	6.95	4.96	1.99	1.40
Carbon dioxide	CO_2	8.75	6.71	2.04	1.30
Hydrogen sulfide	H_2S	8.62	6.53	2.09	1.32
Nitrous oxide	N_2O	8.81	6.76	2.05	1.30
Sulfur dioxide	SO_2	9.70	7.52	2.18	1.29
Acetylene	C_2H_2	9.96	7.91	2.05	1.26
Ethylene	C_2H_4	10.05	7.99	2.06	1.25
Ethane	C_2H_6	11.59	9.51	2.08	1.22
Benzene	C_6H_6	25.4	23.1	2.3	1.10
n-Hexane	C_6H_{14}	31.4	29.1	2.3	1.08

as a unit, but to rotation and vibration as well. If we think simply of a diatomic molecule as a "dumbbell" held together along its line of centers by an elastic spring, then the two atoms may execute vibrations with respect to each other along their line of centers. Further, the molecule as a whole may undergo rotation about axes perpendicular to the line joining the centers of mass of these molecules. These extra motions involve additional terms for the energy of the gas; and if these motions are subject to temperature variation, as they are, additional terms will appear in the heat capacity equation for the gas.

Viscosity of Gases. All fluid bodies, i.e., gases and liquids, possess a property known as the *viscosity*, which may be defined as the resistance that one part of a fluid offers to the flow of another part of the fluid. Viscosity is produced by the shearing effect of moving one layer of the

fluid past another and is quite distinct from intermolecular attraction. It may be thought of as caused by the internal friction of the molecules themselves, and is present in ideal gases as well as in real gases and liquids.

If we think of a fluid as being stratified in layers, then the force f required to move a layer of fluid of area A with a velocity v past another layer a distance ∂ cm away has been shown to be

$$f = \frac{\eta A v}{\partial} \tag{66}$$

Here η is a proportionality constant known as the *viscosity coefficient* of the fluid. The viscosity coefficient may be thought of as the force per unit area required to move a layer of fluid with a velocity of 1 cm per second past another parallel layer 1 cm away. Although the force f may vary with experimental conditions, the viscosity coefficient η is a physical quantity characteristic of each fluid. For gases η is small in magnitude, but in liquids η is fairly large and is of considerable importance in characterizing liquids and their behavior.

It has already been pointed out that according to the kinetic theory of gases the viscosity coefficient of a gas is related to the density d, the mean free path l, and the average velocity v by the relation

$$\eta = \frac{1}{3} v l d \tag{42}$$

Since the mean free path varies inversely as the density of the gas, it may be concluded that the viscosity of an ideal gas should be independent of density, and hence also the pressure. This deduction has been confirmed at relatively low pressures.

Equation (42) may be employed to calculate the mean free path directly from the viscosity coefficients. To do this we need only substitute the value of v from equation (37), in which case l becomes

$$l = \frac{3\,\eta}{vd} = \frac{3\,\eta}{0.921\,d\sqrt{\dfrac{3\,RT}{M}}}$$

$$= \frac{1.88\,\eta}{d\sqrt{\dfrac{RT}{M}}} \tag{67}$$

Once l is thus found it may be inserted into equation (41), and the molecular diameter of the gas molecule σ may be evaluated.

In the cgs (centimeter-gram-second) system of units, the viscosity coefficient of a fluid is expressed in *poises*, a poise being the viscosity

coefficient requiring a force of 1 dyne when A, v, and ∂ are all unity in equation (66). Since this unit is rather large, the viscosities of gases are usually given in *micropoises*, or 10^{-6} poise.

The viscosity of gases generally *increases* with increase in temperature. Thus, $\eta = 5.7$ micropoises for hydrogen at $-258°$ C, while at $300°$ C it is $\eta = 139.2$. Again, whereas the kinetic theory predicts that the viscosity of ideal gases should be pressure independent, actually gases may exhibit considerable variation in viscosity with pressure. For instance, at $35°$ C and atmospheric pressure $\eta = 156$ micropoises for carbon dioxide, but at 80 atmospheres and the same temperature $\eta = 361$ micropoises.

REFERENCES FOR FURTHER READING

1. Hougen and Watson, "Industrial Chemical Calculations," John Wiley & Sons, Inc., New York, 1936.
2. K. Jellinek, "Lehrbuch der physikalischen Chemie," Ferdinand Enke, Stuttgart, 1928, Vol. I.
3. E. Kennard, "Kinetic Theory of Gases," McGraw-Hill Book Company, Inc., New York, 1938.
4. L. Loeb, "Kinetic Theory of Gases," McGraw-Hill Book Company, Inc., New York, 1936.
5. L. Loeb, "Nature of a Gas," John Wiley & Sons, Inc., New York, 1931.
6. D. M. Newitt, "High Pressure Plant and the Properties of Fluids at High Pressures," Oxford University Press, London, 1940.
7. S. F. Pickering, "Relations between Temperatures, Pressures, and Densities of Gases," Circular No. 279, U. S. Bureau of Standards, Washington, D. C., 1925.
8. J. K. Roberts, "Heat and Thermodynamics," Blackie & Son, Ltd., London, 1928.

PROBLEMS

1. At the ice point and under a pressure of 1 m Hg, a given weight of N_2 occupies a volume of 1 liter. At $-100°$ C the same weight of gas under the same pressure occupies a volume of 0.6313 liter. Calculate the absolute zero in degrees centigrade, and give reasons for the observed difference from the accepted value.
2. A gas occupies a volume of 2 liters under a pressure of 720 mm Hg at 25° C. What volume will the gas occupy under standard conditions of temperature and pressure?   *Ans.* 1.736 liters
3. Calculate the density in grams per liter of ammonia at 100° C under a pressure of 800 mm Hg.   *Ans.* 0.585 g/liter
4. Using the simple gas law, find what volume 4 g of CH_4 would occupy at 27° C at a pressure of 2.5 atm.
5. Calculate the pressure exerted by 2 g of O_2 confined in a 2-liter container at 200° C.
6. (a) Find the weight of helium gas necessary to fill a balloon whose capacity is 1,000,000 cu ft, at 1 atm pressure and 27° C. (b) What will be the lifting power of the balloon in pounds per cubic foot under these conditions?
                    Ans. (a) 10,143 lb; (b) 0.0631 lb/cu ft

7. The composition of a mixture of gases in percentage by volume is 30% N_2, 50% CO, 15% H_2, and 5% O_2. Calculate the percentage by weight of each gas in the mixture.

8. Assuming that dry air contains 79% N_2 and 21% O_2 by volume, calculate the density of moist air at 25° C and 1 atm pressure when the relative humidity is 60%. The vapor pressure of water at 25° C is 23.76 mm. *Ans.* 1.171 g/liter

9. Find the total pressure exerted by 2 g of ethane and 3 g of CO_2 contained in a 5-liter vessel at 50° C.

10. At 27° C, 500 cc of H_2, measured under a pressure 400 mm Hg, and 1000 cc of N_2, measured under a pressure of 600 mm Hg, are introduced into an evacuated 2-liter flask. Calculate the resulting pressure.

11. Compare the times of diffusion through a given orifice, and under the same conditions of temperature and pressure, of the gases H_2, NH_3, and CO_2 relative to that of N_2.

12. The time required for a given volume of N_2 to diffuse through an orifice is 35 sec. Calculate the molecular weight of a gas which requires 50 sec to diffuse through the same orifice under identical conditions. *Ans.* 57.15 g/mole

13. What is the total kinetic energy of translation in ergs of 1 mole of a perfect gas at 27° C? In calories?

14. Calculate the root-mean-square velocity in centimeters per second of N_2 molecules at 27° C. Repeat the calculation at 127° C.

15. By means of a mercury vapor pump a vacuum of 10^{-7} mm Hg is obtained within a certain apparatus. Calculate the number of molecules which still remain in 1 cc of the apparatus at 27° C.

16. Calculate the root-mean-square, average, and most probable velocities in centimeters per second of H_2 molecules at 0° C.

17. (a) Using the van der Waals equation, calculate the pressure developed by 100 g of CO_2 contained in a volume of 5 liters at 40° C. (b) Compare this value with that calculated using the simple gas law.
Ans. (a) 11.17 atm; (b) 11.67 atm

18. Using the Beattie-Bridgeman equation explicit in volume, calculate the density in g/cubic centimeter of N_2 at 0° C and 100 atm pressure.

19. At 0° C and under a pressure of 100 atm the compressibility factor of O_2 is 0.927. Calculate the weight of O_2 necessary to fill a gas cylinder of 2 cu ft capacity under the given conditions.

20. The following data were taken in measuring the molecular weight of a certain gas by the Regnault method:

Wt. of evacuated bulb	= 42.5050 g
Wt. of bulb + gas	= 43.3412 g
Wt. of bulb + H_2O	= 365.31 g
Temperature	= 25° C
Pressure (corrected)	= 745 mm

Find the molecular weight of the gas.

21. The elementary analysis of a compound yielded the following results: C, 39.98%; H, 6.72%; and O, 53.30%. In a Victor Meyer determination 0.1510 g of the vaporized compound displaced 33.8 cc of air measured at 25° C over H_2O and at a barometric pressure of 745 mm. Calculate (a) the empirical formula, (b) the approximate molecular weight, and (c) the molecular formula of the compound.
Ans. CH_2O; 115.2 g/mole

22. A sample of vapor weighing 0.180 g occupies a volume of 53.1 cc at 27° C and 760 mm pressure (corrected). The critical pressure of the vapor is 47.7 atm, while the critical temperature is 288.5° C. By use of the Berthelot equation, calculate the molecular weight of the vapor, and compare the result with that calculated by the ideal gas law.

23. The densities of CH_4 at 0° C were measured at several pressures with the following results:

Pressure (atm)	Density (g/liter)
1/4	0.17893
1/2	0.35808
3/4	0.53780

Find the exact molecular weight of CH_4.

24. How much heat will be required to raise the temperature of 10 g of N_2 from 0° C to 100° C at a pressure of 1 atm?

Elementary Principles of Thermodynamics and Their Application to Gases

Introduction. One of the most fundamental manifestations in nature is the energy that accompanies all changes and transformations. Such diversified phenomena as the drop of a stone, the motion of a billiard ball, the impinging of light, the burning of coal, and the growth and reactions of the complex mechanism known as a living being all involve absorption, emission, and redistribution of energy. The most common form in which this energy appears, and the form to which all others tend, is heat. Besides this, there is mechanical energy involved in the motion of all machinery; electrical energy, exhibited by a current in heating a conductor and in doing chemical and mechanical work; radiant energy, inherent in visible light and in radiation in general; and finally, chemical energy, the energy stored in all substances, and which appears when the substances undergo transformation. As diversified and distinct as these various forms may at first glance appear, they are, nevertheless, related to one another, and under certain conditions may be transformed from one into the other. A study of this interrelation of the various forms of energy in a system constitutes the subject of *thermodynamics*.

Since thermodynamic laws deal with energy, they are applicable in general to most phenomena in nature. They hold quite rigidly because they are based on the behavior of macroscopic systems, i.e., systems comparatively large and involving many molecules, rather than on the behavior of microscopic systems in which comparatively few molecules are involved. Moreover, thermodynamics does not consider the time element in transformations; it is interested merely in the initial and final states of a system without any curiosity as to speed with which the change has been accomplished.

Within any system the energy may be kinetic or potential in nature, or both. Kinetic energy is the energy a system possesses by virtue of its motion, be it molecular or motion of the body as a whole. Potential energy, on the other hand, is the energy a system possesses by virtue of its position, i.e., energy due to the structure of the body or due to its

position with respect to other bodies. The total energy content of any system is the sum of the potential and kinetic energies.

The absolute value of the total energy contained in a system cannot at present be determined. The energy differences, however, which appear when a system changes from one state to another can be measured. These differences are expressed in the ordinary work units used in connection with the various forms of energy. Thus the cgs unit of mechanical energy is the *erg*, of electrical energy the *joule*, while of thermal energy the *calorie*. The relation of the unit of mechanical work to the thermal unit is known as the *mechanical equivalent of heat*. The first determinations of this equivalent by Joule laid one of the foundation stones upon which the first law of thermodynamics was reared.

The First Law of Thermodynamics. The first law of thermodynamics is the law of conservation of energy, namely, that *energy can be neither created nor destroyed*. In its present form this law was first enunciated by Mayer in 1842 and by von Helmholtz in 1847. The substance of these two statements is that *in any isolated system the total energy of the system remains constant throughout any operations which may be carried out*. Worded differently, the first law is tantamount to the statement that, for any quantity of a form of energy that disappears, another or other forms of energy will appear in total quantity exactly equal to the amount that disappeared. To be more specific, consider the fate of a quantity of heat q added to a system. This heat will go to raise the internal energy of the system and also to do any outside work the system may perform as a result of the absorption of heat. If we let ΔE be the increase in the internal energy of the system and w be the work done by the system on its surroundings, then by the first law

$$\Delta E + w = q$$

and
$$\Delta E = q - w \qquad (1)$$

Equation (1) is the first law of thermodynamics expressed mathematically. The increase in the internal energy of the system is defined as the difference between the internal energy of the system in its final state E_2 and the internal energy of the system in its initial state E_1. Obviously, therefore, ΔE is given by

$$\Delta E = E_2 - E_1 \qquad (2)$$

and *depends only on the initial and final states of the system*. The values of q, the heat absorbed, and w, the work done by the system, depend, however, on the manner in which the process is conducted. The quantities q, w, and ΔE are directly measurable, but the absolute magnitudes of E_1 and E_2 are at present indeterminate.

When instead of an increase there is a decrease in the internal energy of the system, or when heat is given off by the system, ΔE and q are minus. If work is done on the system by the surroundings, w is negative.

The symbol w represents the total work performed by a system against its surroundings. In a galvanic cell, for instance, w may represent the electrical energy supplied plus, if there is a change in volume, any energy utilized to effect the expansion or contraction against a pressure P. The magnitude of the pressure-volume or mechanical work term is easily derivable as follows:

Consider a cylinder, Fig. 1, of cross-sectional area A, fitted with a frictionless piston. Let the pressure on the piston be P. Then, since pressure is force per unit of area, the total force acting on the piston is $f = P \times A$. If the piston is moved now through a distance dl, the work dw done is

Fig. 1. Pressure-Volume Work

$$dw = fdl = P \times A \times dl$$

But $A \times dl$ is the element of volume dV, swept out by the piston in its motion. Hence,

$$dw = fdl = PdV \qquad (3)$$

and, on integration between the limits V_1 and V_2,

$$w = \int_{V_1}^{V_2} PdV \qquad (4)$$

Equation (4) is perfectly general and applies to the calculation of the P–V work done in any expansion or contraction.

If the only work done by a system is of this type, the expression for the first law becomes

$$\Delta E = q - w$$
$$= q - \int_{V_1}^{V_2} PdV \qquad (5)$$

From equation (5) we see that, when any process is conducted at constant volume, dV, and hence $\int_{V_1}^{V_2} PdV$, is zero, and the equation reduces to

$$\Delta E = q \qquad (6)$$

i.e., all the heat absorbed by a system at constant volume goes to increase the internal energy of the system.
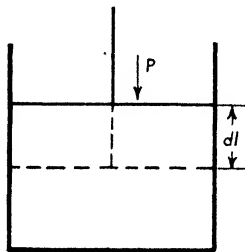
The Heat Content (Enthalpy) of a System. When a process is conducted at constant pressure, however, enough heat must be added not only to take care of the increase in the internal energy, but also to perform any work occasioned by a change in volume. Thermal changes at constant pressure are most conveniently expressed in terms of another function H, called the *heat content* (or frequently, by chemical engineers, *enthalpy*) of a system, which is defined by

$$H = E + PV \tag{7}$$

Since E is characterized by the state of a system only, and since the same is true of the product PV, H, like E, is a function only of the state of the system and is completely independent of the manner in which that state was achieved. Consequently, the change in heat content H may similarly be defined as

$$\Delta H = H_2 - H_1 \tag{8}$$

where H_2 is the heat content of the system in the final state and H_1 the heat content of the system in the initial state. Substituting for H_2 and H_1 their equivalents from equation (7), with the appropriate subscripts, we obtain for ΔH

$$\begin{aligned}
\Delta H &= H_2 - H_1 \\
&= (E_2 + P_2V_2) - (E_1 + P_1V_1) \\
&= (E_2 - E_1) + (P_2V_2 - P_1V_1) \\
&= \Delta E + (P_2V_2 - P_1V_1) \tag{9}
\end{aligned}$$

When the pressure P is constant throughout the process, then

$$\begin{aligned}
\Delta H &= \Delta E + (P_2V_2 - P_1V_1) \\
&= \Delta E + P(V_2 - V_1) \\
&= \Delta E + P\Delta V \tag{10}
\end{aligned}$$

i.e., the change in heat content at constant pressure is equal to the increase in internal energy plus any pressure-volume work done. Hence at constant pressure ΔH represents the heat absorbed by a system in going from an initial to a final state, provided the only work done is P–V work. When the initial and final pressures are not the same, ΔH is calculated not by equation (10) but by equation (9).

Using the notation of G. N. Lewis and M. Randall,[1] heat absorbed, resulting in any way in the increase of a property of a system, is considered positive, while heat evolved, resulting in a decrease in some property of the system, is considered negative. Thus $+ \Delta E$ and $+ \Delta H$ refer to an increase in the internal energy and heat content of a system,

---

[1] " Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.

while $- \Delta E$ and $- \Delta H$ represent a decrease in the same properties of the system.

Heat Capacity. Since no $P-V$ work is done by a system kept at constant volume, any heat added will be utilized to raise the temperature of the system, provided no modifications in form or state occur. The heat absorbed under these conditions per mole of substance for a temperature rise dT, to give an internal energy increase dE, is

$$dE = C_v \, dT \qquad (11)$$

where C_v is the heat capacity at constant volume; it is the amount of heat necessary to raise the temperature of one mole of a substance 1 degree at constant volume. It follows directly from equation (11) that

$$\left(\frac{\partial E}{\partial T}\right)_V = C_v \qquad (12)$$

The subscript v denotes that the change in internal energy takes place at constant volume. Equation (12) is the thermodynamic equation defining the specific heat at constant volume.

When, however, a system absorbs heat not at constant volume but at constant pressure, the heat added goes not only to raise the internal energy of the system but also to perform $P-V$ work. Then

$$dE + P \, dV = C_p \, dT$$
$$[\partial(E + PV)]_P = C_p \, dT \qquad (13)$$

Here C_p is the molar heat capacity at constant pressure. On transposing dT we obtain

$$C_p = \left[\frac{\partial}{\partial T}\left(E + PV\right)\right]_P = \left(\frac{\partial H}{\partial T}\right)_P \qquad (14)$$

which is the thermodynamic definition of C_p.

The Difference $C_p - C_v$ for Ideal Gases. The difference in heat capacities at constant pressure and constant volume for an ideal gas is readily deducible by a thermodynamic argument. The methods employed in this derivation are quite typical of thermodynamics, and illustrate a mode of attack frequently employed. Since $C_p = (\partial H/\partial T)_P$, and $C_v = (\partial E/\partial T)_V$, then

$$C_p - C_v = \left(\frac{\partial H}{\partial T}\right)_P - \left(\frac{\partial E}{\partial T}\right)_V \qquad (15)$$

But, $H = E + PV$. Differentiating this quantity with respect to temperature at constant pressure we obtain

$$\left(\frac{\partial H}{\partial T}\right)_P = \left(\frac{\partial E}{\partial T}\right)_P + P\left(\frac{\partial V}{\partial T}\right)_P \qquad (16)$$

Substituting the value for $(\partial H/\partial T)_P$ from equation (16) in equation (15), the latter becomes

$$C_p - C_v = \left(\frac{\partial E}{\partial T}\right)_P + P\left(\frac{\partial V}{\partial T}\right)_P - \left(\frac{\partial E}{\partial T}\right)_V \qquad (17)$$

The problem now is to relate the first and third terms on the right of equation (17). To do this we proceed as follows. The internal energy E will be, in general, a function of any two of the three variables P, V, T. If we take as our independent variables V and T, then

$$E = f(T, V)$$

and

$$dE = \left(\frac{\partial E}{\partial T}\right)_V dT + \left(\frac{\partial E}{\partial V}\right)_T dV \qquad (18)$$

Dividing both sides of the equation by dT and imposing the condition of constant pressure, we get

$$\left(\frac{\partial E}{\partial T}\right)_P = \left(\frac{\partial E}{\partial T}\right)_V + \left(\frac{\partial E}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P \qquad (19)$$

Equation (19) substituted in equation (17) yields finally

$$C_p - C_v = \left(\frac{\partial E}{\partial T}\right)_V + \left(\frac{\partial E}{\partial V}\right)_T\left(\frac{\partial V}{\partial T}\right)_P + P\left(\frac{\partial V}{\partial T}\right)_P - \left(\frac{\partial E}{\partial T}\right)_V$$

$$= \left(\frac{\partial E}{\partial V}\right)_T\left(\frac{\partial V}{\partial T}\right)_P + P\left(\frac{\partial V}{\partial T}\right)_P \qquad (20)$$

Equation (20) is perfectly general. However, for an ideal gas considerable simplification is possible. First, it will be shown below that for an ideal gas $(\partial E/\partial V)_T = 0$, and hence the first term on the right is zero. Second, since for an ideal gas $PV = RT$, differentiation of the ideal gas law with respect to T at constant P yields

$$PV = RT$$

$$P\left(\frac{\partial V}{\partial T}\right)_P = R$$

and consequently equation (20) becomes

$$C_p - C_v = P\left(\frac{\partial V}{\partial T}\right)_P = R \qquad (21)$$

This is the result obtained previously through the kinetic theory of gases.

The Work of Expansion of Gases. As the work performed in the expansion and contraction of gases is of extreme importance in thermodynamics and physical chemistry, it deserves careful attention. There

are many conditions under which pressure-volume changes may take place; but for all of these the work performed is given in general by the expression

$$w = \int_{V_1}^{V_2} P \, dV \tag{4}$$

However, the manner in which the integration is performed is dependent on the manner in which the volume change is accomplished. For a clear understanding of what follows the student should realize that *the P in the work expression is not necessarily the pressure of the working gas but the pressure against which the gas is working,* i.e., P is the pressure opposing the gas expansion. Once this is realized, no difficulty will be encountered.

According to equation (4), the work done by a gas will be zero when either dV or P is zero. When $dV = 0$, the volume of the system is constant, no expansion takes place, and no work is done. Again, when $P = 0$ the gas in expanding encounters no opposition and hence does no work. The latter process, involving expanding into a vacuum, is called *free expansion.* In both these cases, since $w = 0$, the first law takes the simple form $\Delta E = q$; i.e., any absorbed heat goes merely to increase the internal energy of the gas, or, in other words, to increase its temperature.

When a gas expands against a *constant pressure* from a volume V_1 to a volume V_2, equation (4) integrates simply to

$$\begin{aligned} w &= \int_{V_1}^{V_2} P \, dV \\ &= P(V_2 - V_1) \end{aligned} \tag{22}$$

For $V_2 > V_1$ the process is a positive expansion and work is done by the gas; when, however, $V_1 > V_2$, the process is a contraction (negative expansion), $V_2 - V_1$ and also w are negative, and work must be done on the gas to perform the volume change.

If the pressure against which a gas expands is not constant throughout the process, P must be substituted as a function of V in equation (4) before the expression may be integrated. If the function $P = f(V)$ is not known, but experimental data are available on the variation of the pressure with volume during the process, the integration may be performed graphically by plotting P against V at each point and determining the area under the curve between the volumes V_1 and V_2.

Isothermal and Adiabatic Expansion of Ideal Gases. The energy relations and the state of any gaseous system during the performance of work depend not only on the manner in which work is performed, but also on certain experimental conditions imposed upon the system as a whole. Two such constraints of especial importance are the performance of work under (*a*) isothermal and (*b*) adiabatic conditions.

(a) *Isothermal Expansion of an Ideal Gas.* Any expansion conducted in a system in a manner such that the temperature remains constant during the entire operation is referred to as an *isothermal* expansion. In general a quantity of heat q added to a gaseous system will contribute both to the performance of work and to an increase in the internal energy of the gas. Since the internal energy of an *ideal* gas is a function of temperature only, an increase therein must result in an increase of the temperature of the gas. However, if the temperature of the gas is held constant, the internal energy cannot change, and any external work performed by the gas can be accomplished only at the expense of absorbed heat. This conclusion follows directly from the first law of thermodynamics, for, when E is constant, $\Delta E = 0$, and hence equation (1) becomes

$$q = w \tag{23}$$

We see, therefore, that the criterion for an isothermal process in an ideal gas is that $\Delta E = 0$, and consequently we may write for such a process

$$q = w = \int_{V_1}^{V_2} P dV \tag{24}$$

The magnitude of q obviously will depend on the manner in which the work is performed. If the pressure P against which the ideal gas expands is constant, the work performed, and q, are both given by equation (22), namely,

$$q = w = P(V_2 - V_1) \tag{25}$$

When the pressure is variable, however, the work done, and q, may vary over a wide range up to a maximum value. The conditions for obtaining this maximum work as well as the calculation of it for the case of an ideal gas will be given below.

(b) *Adiabatic Expansion of an Ideal Gas.* Any change in a system taking place under such conditions that heat is neither absorbed nor evolved by the system is said to take place *adiabatically*. The criterion for such a process is that $q = 0$. Making this substitution in equation (1) we get

$$w = -\Delta E \tag{26}$$

From equation (26) it follows that any work in an adiabatic process is done at the expense of the internal energy. As work is performed, the internal energy of the system decreases, and consequently the temperature drops.

An equation that every ideal gas must obey at every stage of an adiabatic expansion can be derived readily from equation (26). Consider

a mole of an ideal gas at a pressure P and a volume V. For an infinitesimal increase in molar volume dV at the pressure P, the work done by the gas is PdV. Since this work is accomplished at the expense of the internal energy of the gas, the internal energy must *decrease* by an amount dE. According to equation (26), therefore,

$$PdV = -dE$$

But from $C_v = (\partial E / \partial T)_V$, $dE = C_v\,dT$. Consequently,

$$PdV = -dE = -C_v\,dT \qquad (27)$$

If the expansion is so controlled that at every stage the external pressure differs from the internal only by an infinitesimal amount, then for P may be substituted $P = RT/V$ from the equation of an ideal gas. Equation (27) becomes thus

$$-C_v\,dT = \frac{RT\,dV}{V}$$

and

$$\frac{C_v}{R}\left(\frac{dT}{T}\right) = \frac{-dV}{V}$$

Considering C_v a constant, and integrating, we have

$$\frac{C_v}{R}\int \frac{dT}{T} = -\int \frac{dV}{V} + \ln C$$

$$\frac{C_v}{R}\ln T = -\ln V + \ln C$$

where $\ln C =$ constant of integration. On rearranging and taking antilogarithms we obtain

$$\frac{C_v}{R}\ln T + \ln V = \ln C$$

$$\ln V T^{C_v/R} = \ln C$$

and

$$V T^{C_v/R} = C \qquad (28)$$

Other forms of equation (28) may be easily derived by eliminating dT or dV instead of P from equation (27). A very common form is one involving P and V, namely,

$$PV^{C_p/C_v} = PV^\gamma = C' \qquad (29)$$

where C' is a constant and $\gamma = C_p/C_v$, the ratio of specific heats. It should be realized that equations (28) and (29) do not displace the ideal gas law $PV = nRT$ but merely *supplement* it. The ideal gas law is applicable under *all* conditions of an ideal gas, while equations (28) and (29) apply only under adiabatic conditions.

The constants C and C' depend on the amounts of gas present and differ from each other numerically. The constants may be eliminated in calculations by taking the ratio of initial to final conditions. Thus

$$P_1 V_1^\gamma = C' \qquad P_2 V_2^\gamma = C'$$

Therefore, $$P_1 V_1^\gamma = P_2 V_2^\gamma$$

and $$\frac{P_1}{P_2} = \left(\frac{V_2}{V_1}\right)^\gamma \qquad (30)$$

Equation (30), like Boyle's law, permits a recalculation of volumes from pressures alone or vice versa. During the adiabatic expansion, however, the temperature of the gas does not remain constant. The initial and final temperatures may be obtained in any instance by substituting the initial and final values of P and V, along with n, in the expression $PV = nRT$.

The use of these equations can be illustrated by an example. Two moles of hydrogen at standard conditions are compressed adiabatically to a volume of 10 liters. For hydrogen, $\gamma = 1.41$. From these data it is desired to find the final pressure and temperature of the gas. The known and unknown quantities are

	Initial	Final
	$P_1 = 1$ atm	$P_2 = ?$
	$V_1 = 2(22.4) = 44.8$ liters	$V_2 = 10$ liters
	$T_1 = 273.1°$ K	$T_2 = ?$
	$n = 2$	$n = 2$
	$\gamma = 1.41$	$\gamma = 1.41$

Applying equation (30), we have for P_2,

$$P_2 = P_1 \left(\frac{V_1}{V_2}\right)^\gamma$$
$$= 1\left(\frac{44.8}{10}\right)^{1.41}$$
$$= 8.3 \text{ atm}$$

Had the expansion taken place *isothermally*, the new pressure would have been 4.5 atm. Knowing now P_2 and V_2, T_2 follows as

$$T_2 = \frac{P_2 V_2}{nR} = \frac{8.3 \times 10}{2 \times 0.0821}$$
$$= 505.5° \text{ K} \quad \text{or} \quad 232.3° \text{ C}$$

The Concept of Reversibility. Unlike the internal energy and heat content, the work obtainable when a system undergoes change depends not only on the initial and final states, but also on how the

change takes place. Conditions can be so arranged that the work performed by the system can vary all the way from zero, in case of expansion into a vacuum, up to the maximum realizable work for a specified transformation when the process is conducted as is to be described now. These facts indicate the great theoretical as well as practical importance of knowing the manner in which the change is accomplished.

Consider a cylinder containing water at its boiling point in equilibrium with its vapor. The cylinder is fitted with a weightless, frictionless piston, and is immersed in a reservoir kept at the same temperature as the water. If the piston is confined by a pressure of 1 atmosphere, this temperature is 100°, and the pressure of the vapor is also 1 atmosphere. Suppose now that the pressure on the piston is diminished by an infinitesimal amount. The piston immediately moves upward, increasing the volume. This change in volume results in a corresponding drop in the vapor pressure of the water. To bring this pressure back to the vapor pressure at the temperature in question, some water evaporates. As this operation is conducted strictly isothermally, the heat necessary for the vaporization of the water is abstracted from the reservoir.

The changes just described, which will proceed as long as the internal pressure of the gas is greater by an infinitesimal amount than the pressure on the piston, will finally result in the complete evaporation of all the water in the cylinder. To stop this evaporation it is necessary merely to increase the confining pressure infinitesimally. This makes the internal and external pressures equal, and neither evaporation of the water nor condensation of the vapor will take place as long as the temperature of the system remains constant.

Consider now, again, what happens when the pressure on the cylinder is increased by an infinitesimal amount. The piston immediately moves downward and compresses the vapor to a pressure above the equilibrium value for the given temperature. To reestablish equilibrium, some of the water condenses. The heat given off by this condensation is taken up by the reservoir, so the isothermal condition of the system is not disturbed.

The essential point about the process described is that *the driving force is at all times only infinitesimally greater than the opposing force* and that by an infinitesimal increase of the opposing force the whole process can be reversed and made to retrace all the stages it has previously occupied. Whenever a process meets these conditions it is said to be *reversible*. Strictly speaking, all naturally occurring processes are *irreversible*. Reversibility can be approached, however. In the potentiometric method for measuring the potentials of galvanic cells, the voltage of the cell is opposed by another voltage until practically no current

flows. By making the opposing voltage only slightly smaller than the voltage of the cell, the cell can be made to discharge; while, if the opposing voltage is increased slightly above that of the cell, the cell can be made to charge. In this manner, any current flowing through the circuit has to do work against a potential which is at all times only slightly less than its own. This arrangement is the closest approximation which can at present be made to a truly reversible process.

Although the idea of reversibility has been developed here for two specific cases, it can be extended to all processes, irrespective of their nature. *Any process*, no matter what it be, *in which the driving force is only infinitesimally greater than the opposing force, and which can be reversed by increasing the opposing force by an infinitesimal amount, constitutes a reversible process.* On the other hand, any change which does *not* meet these requirements is said to be *irreversible*.

Reversibility and Maximum Work. The amount of work a system has to perform to bring about a certain change depends on the opposition the system experiences to the change. The greater that resistance is, the more work must be done by the system to overcome it. To be specific, let us consider the expansion of an ideal gas against a pressure P through an infinitesimal volume change dV. The work done is evidently $P\,dV$. When P is zero, i.e., when the system expands into a vacuum, the piston confining the gas experiences no restraining force, and if it is frictionless and weightless, no work is involved, and $P\,dV = 0$. However, as P is increased from zero, more and more work has to be done as the pressure approaches that of the gas itself. When the latter point is reached, the two forces become balanced and no further change in volume is possible. If we continue to increase the pressure, the pressure on the gas becomes greater than the pressure of the gas, the volume begins to decrease, and work is done *on* rather than *by* the system. From this description it is evident that the work which may be performed by a system is a maximum when the opposing pressure P differs only infinitesimally in magnitude from the internal pressure of the gas itself. But these are exactly the conditions defined for the reversibility of a process. Hence it may be concluded that *maximum work is obtainable from a system when any change taking place in it is entirely reversible.*

The expression for the maximum work performed when an *ideal gas* expands *isothermally and reversibly* may be derived as follows. We have seen that the work resulting from a volume change V_1 to V_2 against a pressure P is given by equation (4), where the indicated integration requires a relation for P in terms of V. Since for a reversible process the external pressure is at all times only infinitesimally lower than the pressure of the gas itself, we may substitute for P in equation (4) the ideal gas pressure $P = nRT/V$. Equation (4) becomes then,

$$w = \int_{V_1}^{V_2} \frac{nRTdV}{V}$$

$$= nRT \int_{V_1}^{V_2} \frac{dV}{V}$$

as the process is also isothermal and T is thus constant. Integrating, we find that

$$w = nRT \ln \frac{V_2}{V_1} \tag{31}$$

which is the expression sought for the *maximum work* obtainable from an isothermal reversible expansion of n moles of an ideal gas from volume V_1 to V_2 at temperature T. Again, since the temperature is constant, $V_2/V_1 = P_1/P_2$ by Boyle's law, and hence equation (31) may also be written in the alternate and equivalent form

$$w = nRT \ln \frac{P_1}{P_2} \tag{32}$$

The application of equation (31), and the distinction between isothermal reversible work and isothermal work against a constant pressure, may best be understood from the following examples.

(*a*) Find the work done when 2 moles of hydrogen expand isothermally from 15 to 50 liters against a *constant pressure* of 1 atmosphere at 25° C. By equation (22),

$$w = P(V_2 - V_1) = 1(50 - 15) = 35 \text{ liter-atm}$$
$$= 35 \times 24.21$$
$$= 847.5 \text{ cal}$$

(*b*) Calculate the work performed when 2 moles of hydrogen expand *isothermally* and *reversibly* at 25° C from 15 to 50 liters. Using equation (31),

$$w = nRT \ln \frac{V_2}{V_1} = 2.3\, nRT \log \frac{V_2}{V_1}$$
$$= (2.3 \times 2 \times 1.99 \times 298) \log \frac{50}{15}$$
$$= 1428 \text{ cal}$$

The heats absorbed during the expansions are equal to w in both cases, and are therefore $q = 847.5$ calories in (*a*), and $q = 1428$ calories in (*b*). The internal energy change is zero in both instances, since there is no change in temperature.

The Maximum Work Function A. We have seen that the amount of work obtained from a process depends on the manner in which the work has been performed. In performing the work reversibly, however, we establish under *isothermal conditions* a definite path of passing from

the initial to the final state along which the maximum work done is definite and dependent only on the two states of the system in question. At constant temperature, then, the maximum work, being a function of the states of the system only, takes on the attributes of such functions as the internal energy E and the heat content H. We may think of a system, therefore, as possessing in each state a certain amount of maximum work content A, such that when the system passes from one state to another, the maximum work performed, w_m, is equal to

$$w_m = -(A_2 - A_1) = -\Delta A \tag{33}$$

A_2 is the maximum work content in the final state, A_1 the work content in the initial state, and $-\Delta A$ the decrease in the maximum work content of the system during the process. The minus sign is included because ΔA is *positive* when the work content of a system *increases*, i.e., *work is done on the system*, while w_m is *positive* when work is done *by* the system. Hence, when all the work in a particular transformation is performed *isothermally and reversibly*, the first law of thermodynamics may be written with the aid of equation (33) as

$$\Delta E = q - w_m = q + \Delta A \tag{34}$$

This equation will be found useful in another chapter.

Maximum Work of Adiabatic Expansion of an Ideal Gas. As an instance of maximum work obtainable under nonisothermal conditions may be given the work performed in the *adiabatic reversible* expansion of an ideal gas. The expression in question is arrived at by combination of equations (4), (29), and $PV = nRT$, and is

$$w_m = \frac{nR(T_2 - T_1)}{(1 - \gamma)} \tag{35}$$

In equation (35) T_1 is the initial temperature of n moles of gas, T_2 is the final temperature, while γ is the ratio C_p/C_v for the gas. Whenever $T_2 > T_1$, w_m is negative and work is done on the gas. On the other hand, when work is done by the gas $T_2 < T_1$, and w_m is positive.

The Joule-Thomson Effect. An ideal gas exhibits no intermolecular attraction, and for it the product PV is a constant at any given temperature at all pressures. Hence when such a gas expands under adiabatic conditions into a vacuum, no heat is absorbed or evolved, no external work or work to separate the molecules has to be performed, and so

$$q = 0, \quad w = 0, \quad \text{and} \quad \Delta E = 0$$

That is, the internal energy of the gas remains constant as well as PV, and consequently the temperature is the same before and after expansion. This is equivalent to saying that at constant temperature the

internal energy of an ideal gas is independent of the volume the gas occupies, or, mathematically,

$$\left(\frac{\partial E}{\partial V}\right)_T = 0 \tag{36}$$

The adiabatic condition was imposed for our purpose merely to prevent any interchange of energy between the surroundings and the gas, and thus to avoid an increase or decrease in the internal energy by heat absorption or evolution.

The situation with *real* gases is different, and was first investigated by Joule and Thomson (Lord Kelvin). Their experimental setup is illustrated schematically in Fig. 2. A tube, thoroughly insulated to approximate adiabatic conditions, was fitted with a porous plug, as indi-
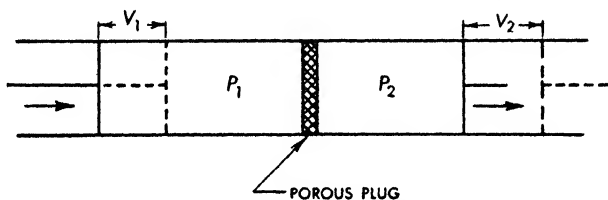
Fig. 2. Joule-Thomson Experiment

cated, to allow gas to be kept on either side of it at the different pressures P_1 and P_2 ($P_1 > P_2$). By applying pressure on the piston on the left slowly enough so as not to change the pressure P_1, a volume of gas V_1 was forced slowly through the porous plug and allowed then to expand to the pressure P_2 and volume V_2 by moving the piston on the right outward. While the expansion was taking place, accurate temperature readings were taken on the gas in the two chambers to ascertain whether the expansion was accompanied by a temperature change.

The work done *on* the system at the left piston is $-P_1V_1$, the work done *by* the system at the right piston is P_2V_2, and hence the net work done *by* the system is

$$w = P_2V_2 - P_1V_1$$

Since the process was conducted adiabatically, $q = 0$, and, therefore,

$$\Delta E = E_2 - E_1 = -w = -(P_2V_2 - P_1V_1)$$

or $$E_2 + P_2V_2 = E_1 + P_1V_1$$

$$H_2 = H_1$$

$$\Delta H = 0 \tag{37}$$

The process was conducted, then, at *constant heat content*. Under these conditions Joule and Thomson observed that all gases, with the excep-

tion of hydrogen, experienced a cooling on expansion, while hydrogen actually became warmer. The extent of the temperature change was found to depend on the initial temperature and pressure of the gas. Later, when helium was discovered, it was shown that this gas as well undergoes a heating effect on expansion. The magnitude of the observed effects may be judged from Table 1, where μ is the *Joule-Thomson coefficient*, defined as

$$\mu = \left(\frac{\partial T}{\partial P}\right)_H \tag{38}$$

It may be thought of as the number of degrees temperature change produced per atmosphere drop in pressure under adiabatic conditions. For a cooling μ is positive, while for an observed heating μ is negative.

TABLE 1

JOULE-THOMSON COEFFICIENTS FOR SEVERAL GASES [1]

Gas	$t°$ C	μ at Pressures (atm)				
		0-6	2	10	15	40
Hydrogen	6.8	− 0.030				
	90.1	− 0.044				
Carbon dioxide	0.0	1.35				1.46
	20.0	1.14	1.21	1.31	1.37	1.20
	40.0	0.96				1.04
	100.0	0.62				.
		$P = 0$	25	50	100	150
Air	− 55	0.44	0.40	0.3	0.28	0.18
	0.6	0.27	0.25	0.24	0.19	0.16
	49	0.20	0.18	0.17	0.15	0.12
	150	0.09	0.09	0.07	0.06	0.05
	250	0.02	0.02	0.02	0.01	0.01

The Inversion Temperature. Hydrogen and helium can be cooled by adiabatic expansion if they are first brought to a sufficiently low temperature. Each gas, before it will cool on expansion, must be below a temperature known as the *inversion temperature*, namely, *the temperature at which the gas exhibits neither cooling nor heating on adiabatic expansion.* Above the inversion temperature gases exhibit a heating effect, while below the inversion temperature a cooling effect is observed. As it happens, all gases except hydrogen and helium are considerably below their inversion points at room temperatures and thus exhibit a positive Joule-

[1] Taylor and Glasstone, "Treatise on Physical Chemistry," third edition, D. Van Nostrand Company, Inc., New York, 1942, Vol. I, p. 57.

Thomson effect. For hydrogen, however, the inversion temperature is at about − 80° C, while for helium the temperature is still lower. Once cooled below their inversion temperatures, hydrogen and helium behave like other gases, cooling on adiabatic expansion.

The Joule-Thomson effect is of great technical and practical importance in the liquefaction of gases. The manner in which it is utilized will be explained in the next chapter where the liquefaction of gases is discussed.

The Second Law of Thermodynamics. Although the first law of thermodynamics establishes the relationship between the heat absorbed and the work performed by a system, it places absolutely no restrictions on the source of this heat and on the direction of its flow. According to the first law there is nothing impossible about a process in which, without any external aid, ice may be used to heat water by extracting heat from the former at a lower temperature and supplying it to the latter at a higher temperature. Yet we know from experience that such a transfer of heat from a lower to a higher temperature will not take place spontaneously. Instead heat is always found to flow of its own accord from the warmer to the colder body; i.e., *the spontaneous flow of heat is always unidirectional from the higher to the lower temperature.*

A similar unidirectionality of change is observed in all natural phenomena. Thus electricity tends to flow only from a point of higher electric potential to one of lower, water will move by itself only from a higher level to a lower, diffusion will occur only from the point of higher concentration to the lower, and all chemical systems under given conditions will tend to undergo reaction in a direction which will lead to the establishment of equilibrium. In fact, all the above observations can be summarized by the statement that *all naturally occurring processes always tend to change spontaneously in a direction which will lead to equilibrium.*

There is still another respect in which the first law is insufficient. In demanding the conservation of energy in all types of transformations, the first law of thermodynamics does not in any way define the ease or extent of convertibility of one form of energy into another. Still, it is an empirical fact that, whereas various forms of energy can be converted readily and completely into heat, the converse process, the conversion of heat into work, can be accomplished only under severely limited conditions. At constant temperature heat can be transformed into work only at the expense of some permanent change in the system involved. For instance, heat may be converted into work by the isothermal reversible expansion of a gas in a cylinder. But to retain this work the gas must remain expanded. If we attempt to return the gas to its original condition, we find that the work obtained in the expansion must be utilized in the compression, and as a result we wind up with the original quantity

of heat and no work. To obtain work from heat by means of a period-ically operating machine, such as a heat engine, it is essential that a temperature drop take place, and that a flow of heat occur from a higher to a lower temperature. Furthermore, even under such conditions not all the heat absorbed by the system can be converted to work, but only a fraction of it determined under ideal conditions by the two tempera-tures between which the operation takes place.

To express the limitations inherent in the convertibility of heat into work, and to indicate the direction of change of all naturally occurring processes, the second law of thermodynamics has been promulgated, largely through the efforts of Clausius and Lord Kelvin. In its general and concise form this law can be stated most conveniently in terms of entropy.[1] As more limited statements may be taken the underlined pas-sages already given in this section, or the following statement by Clausius that *"it is impossible for a self-acting machine, unaided by external agency, to convey heat from a body at one temperature to another body at a higher temperature."* The latter pronouncement merely states that the natural tendency of heat is to flow from the higher temperature to the lower. However, it should be realized that nothing implied in this statement would preclude the possibility of cooling a body below the temperature of its surroundings. Such a cooling can be accomplished, but in order to bring it about work will have to be expended, as in the liquefaction of a gas like ammonia, where the gas has to be compressed before heat can be extracted from it.

·Although the second law was once considered to be of universal valid-ity, certain phenomena observable in conjunction with the Brownian movement have sharply brought to the attention of scientists the fact that the law applies strictly only to systems composed of large numbers of individuals. For instance, in observing in a colloid the random motion of particles which constitutes the Brownian movement, it is found that most of the particles move from the more concentrated portion of the field to the less concentrated, as may be expected. Occasionally, how-ever, just the reverse takes place, and a particle moves toward the more concentrated portion of the field. This is in violation of the second law of thermodynamics. To meet this contingency, the second law is at present considered to be strictly applicable only to statistical systems involving large numbers of particles, but not to the individual units of a system. That this is the correct view is borne out by the connection which can be shown to exist between the second law and the laws of probability governing the behavior of systems containing large numbers of molecules.

The Carnot Cycle. A question of great significance in any considera-tion of energy transformation concerns the extent to which heat is con-

[1] See Chapter **X.**

vertible to other forms of energy which may be utilized to do work. Experience has shown that periodically operating heat engines which absorb a quantity of heat at some temperature T_2 and reject the waste heat at a lower temperature T_1 can convert only a fraction of the absorbed heat into work. Offhand it may be thought that this limited convertibility of heat into work lies in the inefficiency of the engines. However, it can be deduced by theoretical considerations that even an ideal engine, operating under ideal conditions, would be able to convert only a certain fraction of the absorbed heat into work, and that this fraction would be determined only by ⁺he operating temperatures T_2 and T_1, and would be totally independent of the nature of the engine. In other words, there is a natural limitation to the convertibility of heat to work up and above any imperfections which may be present in any contrivances employed.

To establish the above deductions and to arrive at the maximum possible convertibility of heat into work, we must consider the sequence of operations called a *Carnot cycle*. A cycle is any series of operations so carried out that at the end the system is back to its initial state. Any process so conducted is referred to as a complete cyclic process. The one we are at present interested in consists of four distinct steps, two isothermal and two adiabatic, and is so conducted that the net work obtained represents the maximum work that can be derived from a quantity of heat absorbed at one temperature and given out at another, lower, temperature.

Fig. 3. The Carnot Cycle

(1) Isothermal Expansion at T_2. Imagine a cylinder fitted with a weightless, frictionless piston and containing 1 mole of an ideal gas. Let the pressure, volume, and temperature of this gas be P_2, V_2, and T_2 respectively. Allow this gas now to expand *isothermally* and *reversibly* to pressure P_1 and volume V_1. In doing so, the gas absorbs a quantity of heat q_2 equal to the work done. This expansion is represented on the diagram in Fig. 3 by the line $P_2V_2P_1V_1$, and involves the work w_1, given by

$$w_1 = RT_2 \ln \frac{V_1}{V_2} = q_2 \qquad (a)$$
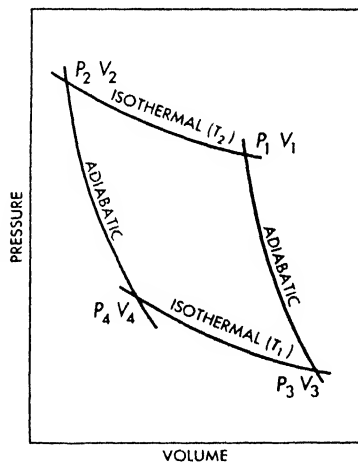
(2) Adiabatic Expansion. Now expand this gas reversibly under adiabatic conditions to a new state P_3V_3. Since the change is adiabatic,

all the work done is at the expense of the internal energy, and involves a drop in temperature from T_2 to T_1. We thus have

$$w_2 = - \Delta E = - C_v(T_1 - T_2) \qquad (b)$$

(3) *Isothermal Compression at T_1.* In the third stage of the cycle the gas is compressed isothermally and reversibly from P_3, V_3, to P_4, V_4 at temperature T_1. During this step a quantity of heat q_1 is given up to the surroundings, and the work performed, w_3, is

$$w_3 = RT_1 \ln \frac{V_4}{V_3} = - q_1 \qquad (c)$$

(4) *Adiabatic Compression.* Finally, the gas is returned to its initial state P_2, V_2, and temperature T_2 by compressing it adiabatically and reversibly. During this stage there is an increase in internal energy equal to the work of compression w_4, or

$$w_4 = - \Delta E = - C_v(T_2 - T_1) = C_v(T_1 - T_2) \qquad (d)$$

The total work performed during this complete cycle is obviously the sum of the work terms involved in the four steps, namely,

$$w = w_1 + w_2 + w_3 + w_4$$
$$= RT_2 \ln \frac{V_1}{V_2} - C_v(T_1 - T_2) + RT_1 \ln \frac{V_4}{V_3} + C_v(T_1 - T_2)$$
$$w = RT_2 \ln \frac{V_1}{V_2} + RT_1 \ln \frac{V_4}{V_3}$$

Consequently,

$$w = RT_2 \ln \frac{V_1}{V_2} + RT_1 \ln \frac{V_4}{V_3} = q_2 - q_1 \qquad (A)$$

Again, since from step (1) $q_2 = RT_2 \ln V_1/V_2$, division of equation (A) by q_2 yields

$$\frac{w}{q_2} = \frac{(q_2 - q_1)}{q_2} = \frac{RT_2 \ln \dfrac{V_1}{V_2} + RT_1 \ln \dfrac{V_4}{V_3}}{RT_2 \ln \dfrac{V_1}{V_2}} \qquad (B)$$

Equation (B) can be simplified considerably. Since the points (P_1, V_1) and (P_3, V_3) lie on the same adiabatic, by equation (28),

$$T_2^{C_v/R} V_1 = C_1 = T_1^{C_v/R} V_3$$

Similarly we have for the points (P_2, V_2) and (P_4, V_4), which lie on the same adiabatic,

$$T_2^{C_v/R} V_2 = C_2 = T_1^{C_v/R} V_4$$

On dividing the first of these equations by the second, we obtain

$$\frac{V_1}{V_2} = \frac{V_3}{V_4}$$

and consequently,

$$\ln \frac{V_4}{V_3} = - \ln \frac{V_1}{V_2}$$

Substituting this value for $\ln V_4/V_3$ in equation (B) we get, finally,

$$\frac{w}{q_2} = \frac{RT_2 \ln \dfrac{V_1}{V_2} - RT_1 \ln \dfrac{V_1}{V_2}}{RT_2 \ln \dfrac{V_1}{V_2}}$$

$$= \frac{T_2 - T_1}{T_2} \tag{39}$$

and

$$w = q_2 \left(\frac{T_2 - T_1}{T_2} \right) \tag{40}$$

In essence equation (40) states that when a system during a reversible cyclical process absorbs a quantity of heat q_2 at temperature T_2, and then undergoes a temperature drop $(T_2 - T_1)$, the external work that may be recovered from the process is equal to the heat absorbed at T_2 multiplied by the ratio $(T_2 - T_1)/T_2$. It is important to observe that the absorption takes place at the higher temperature, and that in accordance with the second law the heat passes from the higher temperature to the lower. Since during this cyclical process all the stages are in every way reversible, *the work done is the maximum possible* under the given temperature conditions. Its magnitude is that indicated by the enclosed area in Fig. 3.

Although equation (40) was derived by using an ideal gas in the cycle described, it is possible to show that the same result can be obtained through the use of any medium, whether ideal or not. Consequently equation (40) is a relation of general validity for the maximum work recoverable from q_2 in any cyclical process operating between the two temperatures T_2 and T_1. This equation may be taken, therefore, as a statement of the optimum, but nevertheless limited, convertibility of heat into work by an engine operating in a cycle between the two temperatures.

The Thermodynamic Efficiency. The ratio w/q_2 in equation (39) is designated the *thermodynamic efficiency* of the process, and represents the maximum fraction of the heat absorbed which is recoverable as work between the temperatures T_2 and T_1. The right-hand side of this equa-

tion points to the startling conclusion that this fraction depends only on the two temperatures between which the cycle is operated, and is independent of all other factors as long as the process is cyclical and reversible. Furthermore, *the thermodynamic efficiency must be the same for all such processes operating under the given temperature conditions.* The necessity for this deduction was pointed out by Carnot, who argued that if any machine were more efficient than one executing a Carnot cycle, the two could be so coupled together as to obtain during a complete cycle a net quantity of work at the higher temperature at the expense of the heat at the lower temperature. But this is contrary to the second law of thermodynamics. Therefore there can be no engine more efficient than a Carnot engine.

No processes in nature approach in efficiency a cyclical reversible operation. Nevertheless, the thermodynamic efficiency, by setting an upper limit, provides a goal toward which improvements in machinery and modes of operation may strive.

The Gibbs-Helmholtz Equation. A very important equation is obtained by eliminating q between the expression for the first law and the one for the work done in the Carnot cycle. The Carnot cycle for a temperature interval $(T_2 - T_1)$ gives for work

$$w = \frac{q_2(T_2 - T_1)}{T_2}$$

For an infinitesimal change in temperature dT, involving an absorption of heat q, the maximum work done, dw, is given by

$$dw = q\,\frac{dT}{T}$$

or

$$q = T\,\frac{dw}{dT} \tag{41}$$

Substituting this value of q in $\Delta E = q - w$, we get

$$\Delta E + w = T\,\frac{dw}{dT} \tag{42}$$

This is the *Gibbs-Helmholtz equation.* In a modified form, using free energy for the maximum work, we shall have occasion to apply it to energy changes involved in chemical reactions. At the present, however, we shall utilize it to derive the Clapeyron equation.

The Clapeyron Equation. Consider the changes involved in the evaporation of unit weight of a substance from liquid to vapor. Let ΔH be the heat of vaporization, V_2 and V_1 the volume of vapor and liquid respectively, and T the temperature at which the transition takes place.

Assume that the liquid is confined in a cylinder under a pressure P, equal to the vapor pressure of the liquid. If a quantity of heat ΔH is absorbed by the cylinder, unit weight of liquid will evaporate. This evaporation involves an expansion against the constant confining pressure P, and the work done is

$$w = P(V_2 - V_1) \qquad (a')$$

The difference between ΔH and the work done must go, then, to increase the internal energy of the system, i.e.,

$$\Delta H - w = \Delta E \qquad (b')$$

Further, differentiation of the work expression (a') with respect to temperature at constant $(V_2 - V_1)$ yields for dw/dT

$$\frac{dw}{dT} = (V_2 - V_1)\frac{dP}{dT} \qquad (c')$$

Substitution of the last two expressions, (b') and (c'), in equation (42) yields then,

$$\Delta E + w = T\frac{dw}{dT}$$

$$\Delta H - w + w = T(V_2 - V_1)\frac{dP}{dT}$$

and
$$\frac{dP}{dT} = \frac{\Delta H}{T(V_2 - V_1)} \qquad (43)$$

This is the *Clapeyron equation*. It is applicable not only to transitions of liquid to vapor, but to any change in form or state of a substance. We may, then, designate ΔH as the heat of transition, T the temperature of transition, V_2 volume of a definite weight of final form, and V_1 volume of the same weight of initial form. dP/dT is the variation of confining pressure with temperature. Illustrations of the application of this equation will be given in connection with the discussion of liquids and solids.

REFERENCES FOR FURTHER READING

1. J. A. V. Butler, "Fundamentals of Chemical Thermodynamics," 2 volumes, Macmillan Company, Ltd., London, 1934–1935.
2. Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.
3. D. J. Martin, "Introduction to Thermodynamics for Chemists," Longmans, Green and Company, New York, 1933.
4. Kiefer and Stuart, "Principles of Engineering Thermodynamics," John Wiley & Sons, Inc., New York, 1930.
5. C. N. Hinshelwood, "Thermodynamics for Students of Chemistry," E. P. Dutton & Company, Inc., New York, 1926.
6. M. Planck, "Treatise on Thermodynamics," Longmans, Green and Company, New York, 1927.

PROBLEMS

1. A weight of 1 kg falls freely to a platform from a height of 10 m. What amount of heat in calories will be evolved when the weight strikes the platform? What work is done?

2. A current of 1 amp flows through a 100-ohm coil. Find the amount of heat in calories dissipated per minute.

3. A piston whose area is 6 sq in. moves through a distance of 8 in. against a constant pressure of 3 atm. Calculate the work done (a) in joules and (b) in calories.

4. Calculate the minimum amount of work necessary to compress 20 g of O_2 from a volume of 10 to a volume of 5 liters at $0°$ C. Is heat absorbed or evolved in the process, and how much?   *Ans.* $w = q = -117.5$ cal

5. A gas in expanding against a constant pressure of 1 atm from 10 to 16 liters absorbs 30 cal of heat. What is the change in internal energy of the gas?

6. Two liters of N_2 at $0°$ C and 5 atm pressure are expanded isothermally and reversibly until the confining pressure is one atmosphere. Calculate the amount of work done in joules and the heat absorbed in calories.
           Ans. $w = 1630.1$ joules; $q = 389.7$ cal

7. What would be the work done and the heat absorbed if in the preceding problem the N_2 under the initial conditions were expanded against a constant pressure of 1 atm? Compare your answers with the results of the preceding problem.

8. Calculate the work done by 10 g of CO_2 in expanding isothermally and reversibly from a volume of 5 liters to 10 liters at $27°$ C. What are q, ΔE, and ΔH for the process? Assume CO_2 to be a perfect gas.
           Ans. $w = q = 93.9$ cal; $\Delta H = \Delta E = 0$

9. Using the van der Waals equation, find the minimum work necessary to compress 1 mole of CO_2 from a volume of 10 liters to 1 liter at $27°$ C. Compare this value with the result which would be obtained if CO_2 behaved as a perfect gas.

10. Eight grams of O_2 at $27°$ C and under a pressure of 10 atm are permitted to expand adiabatically and reversibly until the final pressure is 1 atm. Find the final temperature and the work done in the process. Assume that $C_p = 7/2\,R$ for O_2.   *Ans.* $t = -117.7°$ C; $w = 179.6$ cal

11. Ten grams of N_2 at $17°$ C are compressed adiabatically and reversibly from 8 to 5 liters. Calculate the final temperature and the work done on the gas. What are ΔE and ΔH for the process? Assume $C_p = 7/2\,R$.

12. It is desired to lower by adiabatic expansion the temperature of NH_3 from $27°$ C at 1 atm to $-3°$ C. What will have to be the final pressure? For NH_3 $C_p = 8.9$ calories per mole.

13. A certain fuel furnishes 7000 cal/gram. Calculate the maximum work obtainable per gram in a heat engine in which (a) Hg at its normal boiling point, and (b) H_2O at its normal boiling point, are used in the boiler. Assume in each case that the condenser is at $40°$ C.

14. Compare the maximum thermal efficiency of a steam engine in which steam under a pressure of 2 atm is used in the boiler with one in which steam under 1 atm is used. In each case the condenser is maintained at $40°$.

15. How much work is done in condensing 1 lb of steam at its boiling point under a constant pressure of 1 atm? What are q and ΔE for the process?
           Ans. $w = -19,240$ cal; $q = -244,600$ cal; $\Delta E = -225,360$ cal

Liquids

General Characteristics of Liquids. Although liquids along with gases are classified as fluids, they differ from gases in several essential respects. Liquids, unlike gases, are characterized by a definite volume which is independent of their container; and, compared with gases, they are only slightly compressible. The densities and viscosities of liquids are greater than those of gases. All gases will mix in all proportions, but certain liquids are partially and sometimes completely immiscible. Thus, alcohol and water mix in all proportions, ether and water only in certain limited proportions, and mercury and water not at all.

From the standpoint of kinetic theory, a liquid may be considered as a continuation of the gas phase into the region of small volumes and very high molecular attractions. The cohesive forces in a liquid must be stronger than those in a gas at even high pressures, for they are high enough to keep the molecules confined to a definite volume. Still, the molecules within the liquid must not be thought of as rigidly fixed. They have some freedom of motion, but this motion is considerably restricted, and hence the mean free path is much shorter than in the gas phase.

At best, our knowledge of the nature of the liquid state is still very incomplete. Because of the proximity of molecules to each other within the liquid, effects frequently manifest themselves in liquids which, if present, are of only secondary significance in gases. Thus, because of purely molecular (van der Waals) forces, and in certain cases electrical (dipole) forces, many liquids exhibit a tendency to cluster or associate, and also to orient themselves in some definite manner. The situation within a liquid is very complex, and the progress made in unraveling the multitudinous effects has been rather slow. It can be truthfully said that we still do not understand well our commonest liquid, water.

Critical Phenomena in Liquids. If a liquid, such as water, is sealed in an evacuated tube, a certain amount will evaporate to form vapor. This vapor will exert a pressure just as any gas does, and, provided the temperature is maintained constant, an equilibrium will be established between the liquid and vapor phases. The vapor pressure established is characteristic for each liquid and is a constant at any given

temperature; it is known as the *saturated vapor pressure* of the liquid. The saturated vapor pressure increases continuously with temperature. Thus, at 25° C the vapor pressure of water is 23.76 mm Hg, while at 100° C it is 760 mm Hg. As the water in the sealed tube is heated further, more and more water evaporates and the pressure continues to increase. At all times there is a definite line of demarcation, or meniscus, between the liquid and vapor phases. When we reach the temperature of 374° C, however, the meniscus becomes indefinite, fades into the vapor, and disappears. At this temperature the physical properties of liquid and vapor become identical, and no distinction can be observed between the two. A liquid in this condition is said to be at the *critical point*. The temperature, saturated vapor pressure, and specific volume corresponding to this point are designated the *critical temperature, critical pressure*, and *critical volume* respectively. Their values, which are constant and characteristic for each substance, are known as the *critical constants*. For water the critical constants are: $t_c = 374°$ C, $P_c = 217.7$ atmospheres, and $V_c = 45$ cc per mole.

On heating the sealed tube even slightly above the critical temperature, no evidence can be found of the presence of liquid. The whole mass is gaseous and remains in that state no matter how high it is heated, or how large an external pressure is applied. Since the phenomena described for water are exhibited by all liquids, it must be concluded that *no liquid can exist as such at temperatures above the critical under any applied pressure.*

The critical phenomena are reversible. When the gas in the sealed tube is cooled below the critical temperature, if the pressure is sufficiently high the meniscus reappears, and again we have the two phases, liquid and vapor.

The *P–V–T* Relations of Gases and Liquids. The first complete data on the *P–V–T* relations of a substance in both gaseous and liquid states were obtained by Andrews [1] on carbon dioxide. In his experiments the carbon dioxide was confined by mercury in a graduated capillary tube sealed to a copper jacket filled with water. By applying pressure on the water, pressures up to 400 atmospheres could be built up within the gas. The volume of gas or liquid corresponding to any given temperature and pressure could be read on the calibrated capillary tube. In this manner Andrews was able to obtain a complete record of the variation of the volume of carbon dioxide with pressure at any given temperature, and to show that the critical temperature of carbon dioxide is 31° C at a critical pressure of 73 atmospheres.

Figure 1 shows the plot of pressure vs. volume for carbon dioxide at various constant temperatures. Each constant *P–V* plot is called an

[1] Andrews, Trans. Roy. Soc., **159**, 583 (1869).

isothermal. The data on which the plot is based are not due to Andrews, but are the composite data of several subsequent investigators as compiled in the International Critical Tables. The 48.1° C isothermal is very similar to the hyperbolic plot demanded by Boyle's law, and shows no presence of liquid carbon dioxide even at the highest pressures
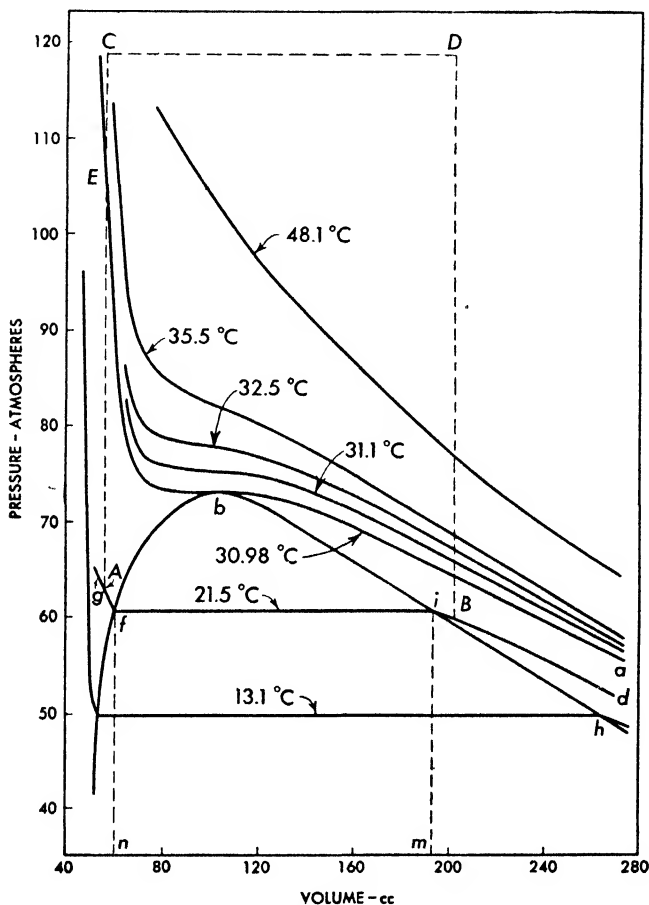
Fig. 1. Isothermals of CO_2

attained. The same conditions obtain at 35.5° C, 32.5° C, and 31.10° C, except that now the data indicate that Boyle's law when applied to carbon dioxide is considerably in error, since the gas does not behave ideally. At 30.98° C, however, the carbon dioxide remains gaseous only up to a pressure of 73 atmospheres (line *ab*). At 73 atmospheres (point *b*) liquid first appears, and since this is the highest temperature at which liquid is observed, 30.98° C must be the critical temperature of carbon

dioxide. Further increase in pressure at this temperature (line bE) shows only the presence of liquid, and consequently this line must represent the compressibility of liquid carbon dioxide at this temperature. Below 30.98° C, the behavior of the gas on compression is quite different, as may be judged from the 21.5° C and 13.1° C isotherms. At 21.5° C, for instance, only gas exists on compression along line di. At i liquid, of specific volume n, first appears, and the pressure of the system remains constant thereafter as long as both gas and liquid are present. At this stage further application of pressure results merely in further condensation of gas until point f is reached. At f all the gas has been condensed, and further application of pressure results merely in compression of the liquid, as is shown by the steep line fg. At lower temperatures the behavior is similar to that at 21.5° C, except that the horizontal portions, corresponding to the range of coexistence of liquid and vapor, become longer the lower the temperature.

It may be concluded from this explanation that, in the area to the left of the dome-shaped area and below the line bE, only liquid carbon dioxide will exist; to the right of the line bE and to the right of the dome-shaped area, only gaseous carbon dioxide will exist; while within the dome-shaped area is the range of coexistence of liquid and vapor carbon dioxide.

All gases upon isothermal compression behave similarly to carbon dioxide. For each, of course, the curves will be displaced in line with the characteristics and critical temperature of the gas in question. Thus, for example, the critical temperature of helium is − 268° C, and the dome-shaped area is moved downward, while for chlorine the critical temperature is 144° C, and the dome-shaped area is moved above that for carbon dioxide.

The Principle of Continuity of States. For further theoretical considerations it is essential to show that the liquid state does not represent a sharp and discontinuous transition from the gaseous state, but is rather a continuation of the gaseous phase into the region of very strong intermolecular attractions and small volumes. This can be shown from the following considerations. Suppose we wish to convert liquid carbon dioxide at 21.5° C and the pressure given by point A in Fig. 1 to gaseous carbon dioxide at the same temperature and the pressure given by point B. The most obvious way to accomplish this transformation is to follow the 21.5° C isotherm and reduce the pressure along $AfiB$. In doing this gas appears suddenly and discontinuously, and coexists with liquid along fi until finally all liquid disappears at i. The same transformation may, however, be accomplished in another way. If the liquid at A is heated at constant volume, increase of temperature will lead to increased pressure, and the mass will move along the line AEC. As long as the carbon

dioxide is below the critical isotherm, point E, the carbon dioxide is liquid; as soon as the carbon dioxide passes the critical isotherm, however, it becomes gaseous. At the critical temperature, as we have seen, the liquid passes to gas imperceptibly and continuously, and hence in heating the liquid from A to C we convert it without discontinuity to gas. The gas at C may now be expanded to D at constant pressure by heating, and then compressed again at constant volume from D to B. By this series of operations we can convert liquid to gaseous carbon dioxide at 21.5° C without introducing any discontinuity between the phases.

The implication involved in this principle of the continuity of the gaseous and liquid states is highly important. It suggests that if we have an equation of state which is satisfactory in the region of high pressures and low temperatures, that equation should be applicable also to the conditions prevailing at the critical point and to the liquid itself. We shall see now how the van der Waals equation meets these requirements.

Application of van der Waals' Equation to the Isothermals of Carbon Dioxide. By substituting $n = 1$ and the values of the constants a and b for carbon dioxide in van der Waals' equation, namely,

$$\left(P + \frac{a}{V^2}\right)(V - b) = RT$$

we can calculate for any given temperature the P–V relationships above, at, and below the critical temperature. The results of such a calculation are summarized in Fig. 2. The plot is, in general, similar to the one obtained experimentally by Andrews. At t_1, for instance, which is above the critical temperature, the P–V relationship is hyperbolic and corresponds closely to that of the 48.1° C isotherm in Fig. 1. At t_c, which is the critical temperature, a slight break is observed at a, the critical point, which is again in accord with observation. However, below the critical temperature, the range determining the coexistence of liquid and gas is indicated by a continuous S-shaped portion as bcd at t_3, rather than by the horizontal constant pressure range actually observed. In this respect, therefore, and in point of strict quantitative agreement with observed data, the van der Waals equation leaves something to be desired. Nevertheless, some investigators have found that by compressing the gas very carefully part of the curve bc may be realized, though only in an unstable condition. Similarly if the pressure on a liquid be released slowly, part of curve cd can be obtained, but again the condition is unstable.

Determination of van der Waals' Constants a and b. If it be assumed that van der Waals' equation is applicable at the critical point, then the van der Waals constants a and b for any gas can be calculated

from the critical constants of the gas in the following manner. On expanding and rearranging the equation we have

$$\left(P + \frac{a}{V^2}\right)(V - b) = RT$$

$$PV^3 - V^2(RT + Pb) + aV - ab = 0$$

and

$$V^3 - \left(\frac{RT + Pb}{P}\right)V^2 + \left(\frac{a}{P}\right)V - \left(\frac{ab}{P}\right) = 0 \qquad (1)$$

This is a cubic equation in V, and for any given value of P and T will yield three separate solutions for V. The three roots of this equation may all be real, or one may be real and positive and the other two imaginary. Thus in Fig. 2 the equation yields the three roots d, c, and
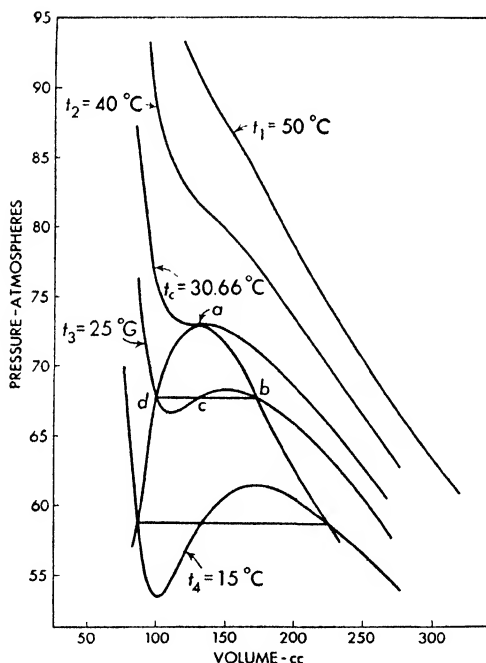
Fig. 2. Isothermals of CO_2 according to van der Waals' Equation

b at t_3, while at t_1 it yields only one real root. However, at the critical point a the three roots are not only real and positive, but also identical and equal to V_c. Hence the difference $(V - V_c) = 0$, and consequently,

$$(V - V_c)^3 = 0 \qquad (2)$$

On expansion by the binomial theorem equation (2) becomes

$$V^3 - (3 V_c)V^2 + (3 V_c^2)V - V_c^3 = 0 \qquad (3)$$

At the critical point equations (3) and (1) must be identical. On comparing and equating coefficients we get

$$3 V_c = \frac{RT_c + bP_c}{P_c} \tag{4}$$

$$3 V_c^2 = \frac{a}{P_c} \tag{5}$$

$$V_c^3 = \frac{ab}{P_c} \tag{6}$$

From equation (5) a follows as

$$a = 3 V_c^2 P_c \tag{7}$$

while from equations (5) and (6) b is given by

$$b = \frac{V_c}{3} \tag{8}$$

Thus a and b may be calculated from known values of P_c and V_c, or vice versa.

Usually V_c is the critical constant known least accurately, and it is therefore preferable to calculate a and b from T_c and P_c only. This can readily be done. On eliminating V_c between equations (4) and (8) we get

$$b = \frac{RT_c}{8 P_c} \tag{9}$$

Again, on combining equations (4), (8), and (5), a follows as

$$a = \frac{27}{64} \frac{R^2 T_c^2}{P_c} \tag{10}$$

In making calculations involving these equations, the units employed must correspond for all terms involved in the equations.

Combination of equations (4) and (8) leads also to the value of R in terms of the critical constants, namely,

$$R = \frac{8}{3} \frac{P_c V_c}{T_c} = 2.67 \frac{P_c V_c}{T_c} \tag{11}$$

Although the van der Waals equation predicts the coefficient in equation (11) to be 2.67, the values for it calculated from experimental data are generally higher and differ for various gases. Thus for helium this constant comes out to be 3.18, while for water it is 4.97. These differences are due to inaccuracies inherent in the van der Waals equation.

The Critical Constants of Gases. Table 1 gives the critical constants of a number of gases. Instead of the critical volume is given the critical density, which is the weight of substance at the critical point per

TABLE 1

CRITICAL CONSTANTS OF GASES

Gas	t_c (° C)	P_c (atm)	d_c (g/cc)	Boiling Point (° C)	$\dfrac{T_c}{T_b}$
Ammonia	132.4	111.5	0.235	− 33.4	1.69
Argon	− 122	48	0.531	− 185.7	1.73
Carbon dioxide	30.98	73.0	0.460	− 78.5 (sub)	—
Carbon monoxide	− 139	35	0.311	− 192	1.65
Chlorine	144.0	76.1	0.573	− 34.6	1.75
Ethane	32.1	48.8	0.21	− 89.3	1.66
Ethyl alcohol	243.1	63.1	0.2755	78.4	1.47
Ethylene	9.7	50.9	0.22	− 103.9	1.67
Helium	− 267.9	2.26	0.0693	− 268.9	1.24
Hydrogen	− 239.9	12.8	0.0310	− 252.7	1.63
Neon	− 228.7	25.9	0.484	− 245.9	1.63
Nitric oxide	− 94	65	0.52	− 151	1.47
Nitrogen	− 147.1	33.5	0.3110	− 195.8	1.63
Oxygen	− 118.8	49.7	0.430	− 183	1.71
Propane	96.81	42.01	0.226	− 44.5	1.62
Toluene	320.6	41.6	0.292	110.8	1.54
Water	374.0	217.72	0.4	100.0	1.74

cubic centimeter. The critical volume is obtained by dividing the molecular weight of the substance by the critical density.

Several empirical rules have been proposed which either facilitate the determination or permit an estimate of some of the critical constants of a substance. One of these is the rule of Cailletet and Mathias or the law of the rectilinear diameter. Cailletet and Mathias found that when the mean values of the sum of the densities of liquid and saturated vapor of a substance are plotted against the temperature, the plot is a straight line. This is shown in Fig. 3. The equation of the line is, of course,

$$T = A + B\left(\frac{d_l + d_v}{2}\right) \qquad (12)$$

where d_l is the density of the liquid at any temperature T, d_v the density of the saturated vapor at the same temperature,
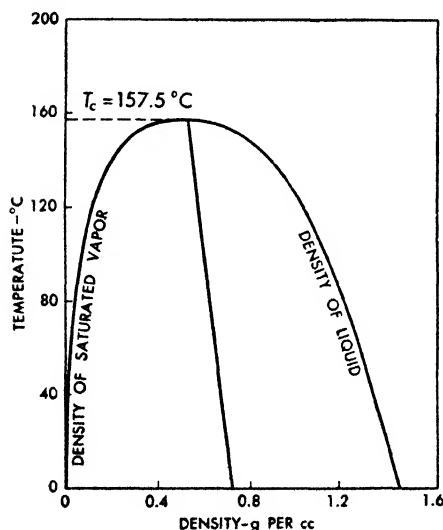
Fig. 3. Linear Variation of Mean Density of SO_2 with Temperature

and A and B constants evaluated from the plot. Once the equation is determined, the critical density may be calculated with ease, for at the critical temperature T_c, $d_v = d_l = d_c$ and the equation reduces to

$$T_c = A + B\left(\frac{d_l + d_v}{2}\right) = A + B\left(\frac{2\,d_c}{2}\right) = A + Bd_c \qquad (13)$$

Substitution of T_c yields then the critical density. The critical densities can usually be obtained more accurately in this manner than by direct measurement at the critical point.

Guldberg and Guye have suggested that the ratio of the critical temperature to the boiling point T_b, both on the absolute scale, is 1.5, or

$$\frac{T_c}{T_b} = 1.5 \qquad (14)$$

That this relation is only approximate may be judged from column 6 of Table 1, where the ratios vary from 1.24 for helium to 1.75 for chlorine. However, where direct measurements of the critical temperature of a substance are not available, the relationship is occasionally useful in estimating T_c from the boiling point.

The Principle of Corresponding States. If we substitute in the van der Waals equation the values of a, b, and R as given by equations (9), (10), and (11), we obtain

$$\left(P + \frac{3\,V_c^2 P_c}{V^2}\right)\left(V - \frac{V_c}{3}\right) = \frac{8}{3}\frac{P_c V_c T}{T_c} \qquad (15)$$

Dividing both sides of equation (15) by $P_c V_c$, we get

$$\left(\frac{P}{P_c} + \frac{3\,V_c^2}{V^2}\right)\left(\frac{V}{V_c} - \frac{1}{3}\right) = \frac{8}{3}\frac{T}{T_c}$$

or,

$$\left(P_r + \frac{3}{V_r^2}\right)(3\,V_r - 1) = 8\,T_r \qquad (16)$$

where

$$P_r = \frac{P}{P_c}$$

$$V_r = \frac{V}{V_c}$$

$$T_r = \frac{T}{T_c}$$

P_r is termed the *reduced pressure*, V_r the *reduced volume*, and T_r the *reduced temperature*. Expressed in terms of P_r, V_r, and T_r, equation (16) involves no constants characterizing the individuality of various substances and should therefore be generally applicable to all liquids and gases. It is known as a *reduced equation of state*. Its physical meaning is that at any given corresponding temperature and pressure, i.e., at any given value of T_r and P_r, all liquids and gases should have the same corresponding volumes, V_r.

The principle of corresponding states is only approximately correct, but it does suggest that frequently better correlation of experimental data may be obtained when the various substances are in corresponding states, i.e., at equal values of T_r, V_r, or P_r. The principle finds frequent and useful application in thermodynamic and chemical engineering calculations, especially at elevated pressures. For examples see Hougen and Watson,[1] Watson and Smith,[2] and Newton.[3]

Liquefaction of Gases. The particular method employed in the liquefaction of a gas depends on the nature of the gas. Vapors of substances which are liquid at or near room temperature and atmospheric pressure are condensed simply by cooling. Other substances which are liquid at lower temperatures may be condensed either by pressure or by a combination of cooling and compression. Such gases as chlorine, sulfur dioxide, hydrogen sulfide, ammonia, methyl chloride, and "freon" (CF_2Cl_2) may be liquefied readily by compression alone. Cooling reduces considerably the pressure required for liquefaction, as may be seen from Fig. 1. With the "permanent" gases, however, such as oxygen, nitrogen, hydrogen, and helium, application of pressure alone will not produce liquefaction, and more involved methods of cooling, compression, and even expansion are required before the gases will liquefy.

Any gas, before liquefaction is possible, must be cooled below its critical temperature. Since their critical temperatures are very low, as may be seen from Table 1, liquefaction of the "permanent" gases requires intense cooling as well as considerable compression. To attain these low temperatures two general principles, or a combination of the two, are employed, namely, (1) adiabatic expansion, in which advantage is taken of the Joule-Thomson effect to attain cooling; and (2) allowing the gas to cool itself by performing work in an adiabatic expansion against a piston. These two methods are exemplified in the Linde and Claude processes for the liquefaction of air.

The basic principle of the Linde process is the adiabatic Joule-Thomson expansion and consequent cooling of the air. The steps in the process

[1] Hougen and Watson, "Industrial Chemical Calculations," second edition, John Wiley & Sons, Inc., New York, 1936, pp. 396–400.

[2] Watson and Smith, National Petroleum News, July 1, 1936.

[3] Newton, Ind. Eng. Chem., **27**, 302 (1935).

are in outline as shown in Fig. 4. Air is first compressed to approximately 100 atmospheres. During the compression most of the water in the air condenses and is removed. The heat generated in compression is removed by passing the gas through coils C, refrigerated by water or ammonia. The dry gas is passed, then, through a copper spiral coil S, from which it is expanded to almost at-
mospheric pressure through a controlled valve V. The issuing gas, cooled now due to the Joule-Thomson effect, passes over the copper coil and cools further the incoming compressed gas. On repeating the cycle several times, the temperature of the expanding gas finally drops far enough to
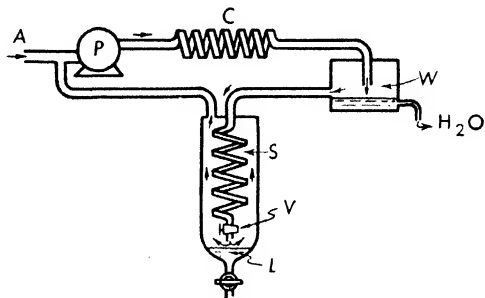
Fig. 4. Linde Process for Liquefaction of Air

condense part of the air to liquid, which collects in the bottom of the chamber L and can be drawn off. Any uncondensed air is recirculated.

In the Claude process the gas, instead of being permitted to expand freely, is forced to do work against a confining piston. Since the gas is adiabatically insulated, work is achieved at the expense of the internal energy of the gas, and a cooling results. The various stages of the process can be understood from Fig. 5 and the following description. The air,
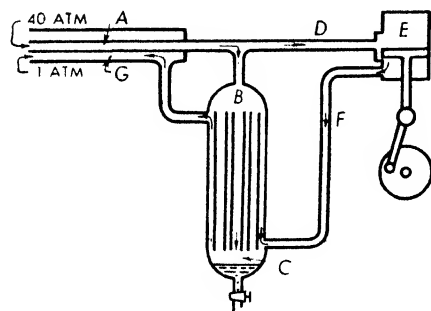
Fig. 5. Claude Process for Liquefaction of Air

freed of carbon dioxide and water vapor, is compressed to about 40 atmospheres and cooled by passing through pipes over which water flows. The air A, entering the apparatus, is subdivided now into two streams, one entering the condenser B directly, while the other, D, is led to the cylinder E, where it expands adiabatically to do work against the piston. The gas thus cooled leaves the cylinder at F, and is

utilized for further refrigeration by being passed over the outside of tubes B and into the precooler tube G. The air entering at B is thus cooled sufficiently to condense, and collects at C where it may be drawn off.

The Claude process is more economical than the Linde process in that it permits the recovery of part of the work expended in compression,

work which is lost completely in the free expansion of the Linde process. The work thus gained may be utilized to operate the compressors.

Easily liquefiable gases are at present of large technical and practical importance in household refrigeration and air conditioning. The gases most commonly used for this purpose are sulfur dioxide, ammonia, methyl chloride, and dichloro-difluoromethane or "freon." In the laboratory the other refrigerants most frequently employed are ice, liquid air, and a mixture of "dry ice" (solid carbon dioxide) and alcohol, ether, or acetone. With one of the last named mixtures temperatures of -80 to $-90°$ C can be obtained. Liquid air will give a temperature of $-180°$ C, while, if needed, liquid hydrogen can give a temperature of $-250°$ C.

The Vapor Pressure of Liquids. Attention has already been called to the fact that a liquid when placed in a container will partially evaporate to establish a pressure of vapor above the liquid. The pressure established will depend on the temperature of the liquid, and will be a constant at any given temperature. The phenomenon is easily explainable in terms of the kinetic theory. The molecules within the liquid, like those in a gas, do not all have the same velocity at a given temperature, but range in velocities from zero for some to very high velocities for others. Most of them, however, are grouped about a mean velocity which is determined by the temperature. When certain molecules within the liquid move in the direction of the surface, and possess sufficient kinetic energy to overcome the large attractive forces of the rest of the molecules, they will escape from the surface into the space above and establish there a vapor pressure. The process does not proceed indefinitely, for some molecules in the gas phase will collide with the surface and will be recaptured by the liquid. When the rate of escape from the liquid becomes equal to the rate of return of molecules to the liquid, an equilibrium is established, and the pressure then remains constant as long as the temperature remains unchanged. This constant vapor pressure is referred to as the *saturated vapor pressure* of the liquid at the particular temperature. As long as this vapor pressure is maintained, the liquid exhibits no further tendency to evaporate. At any lower pressure the liquid will evaporate into the gas phase, while at any higher pressure vapor will tend to condense until the equilibrium pressure is reestablished.

Inasmuch as the evaporation involves the escape of molecules of high kinetic energy, the average kinetic energy of the molecules within the liquid must decrease, and hence the temperature of the liquid tends to drop. To preserve isothermal conditions, heat must be supplied. The quantity of heat required depends on the liquid in question, the amount evaporated, and the temperature. For any liquid at a given temperature

the amount of heat required per unit weight of liquid is a definite quantity known as the *heat of vaporization* of the liquid. It is the difference in the heat contents of vapor and liquid respectively, or

$$\Delta H_v = H_v - H_l \tag{17}$$

where ΔH_v is the heat of vaporization, H_v the heat content of vapor, and H_l the heat content of liquid. For an evaporation ΔH_v is always positive, i.e., heat is always absorbed, while for a condensation ΔH_v is always negative and equal numerically to the heat absorbed in the vaporization. As may be expected from the definition of the heat content, ΔH_v represents the sum of the difference in the internal energy of vapor and liquid, $\Delta E_v = E_v - E_l$, and the work involved in the expansion from liquid to vapor; namely,

$$\Delta H_v = \Delta E_v + P\Delta V \tag{18}$$

where P is the vapor pressure and $\Delta V = V_v - V_l$.

Measurement of Vapor Pressure. The various methods available for measuring the vapor pressure of a liquid may be classified generally into static and dynamic methods. In the static methods the liquid is permitted to establish its vapor pressure without being disturbed in any way, while in the dynamic methods the liquid either is boiled or has a stream of inert gas passing through it. The line of demarcation between these two classifications is not always sharp, and a particular procedure may actually be a combination of the two.

The most obvious method of measuring the vapor pressure of a liquid is to introduce a small quantity of liquid above the mercury surface of a barometer, and to observe the depression in the barometric reading caused by the pressure of vapor above the mercury. This method, however, does not lend itself to precise measurements, for temperature control is difficult, and at elevated temperatures the mercury itself exhibits a significant vapor pressure. Again, any air present in the liquid cannot be removed conveniently, and will tend to give high values for the vapor pressure.

The Isoteniscope Method of Menzies and Smith is much more precise, flexible, and convenient for the measurement of the vapor pressures of a substance over a range of temperatures. A simple laboratory setup is illustrated in Fig. 6. The isoteniscope bulb B is filled one-half to three-quarters full with the liquid to be studied, and the U-shaped portion of the tube, C, is filled to a depth of 2 or 3 cm with the same liquid. The isoteniscope is then attached to the rest of the apparatus and surrounded by a water bath A, whose temperature is measured by thermometer D. E is a barometric leg for measuring the pressure in the apparatus, while G is a large bottle to smooth out pressure fluctuations in the system.

This bottle can be connected alternately to a suction pump or the air. In operation the system is evacuated until the liquid boils vigorously at B to expel all air from BC. The bath A is then adjusted to the desired temperature, and air is slowly admitted to the system until the liquid levels in the U-tube C are exactly equal. Under these conditions the pressure on either side of the U-tube must be the same. Hence, the
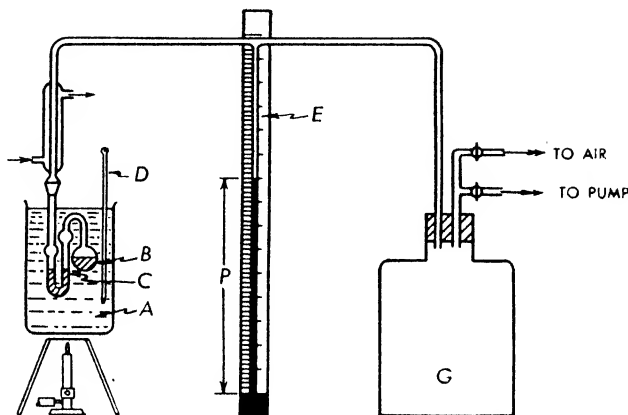
Fig. 6. Isoteniscope Assembly for Determination of Vapor Pressure

vapor pressure in B must be the same as the pressure in the rest of the apparatus, and can be obtained from the reading of the barometer and the mercury column at E. The difference between the barometric pressure and that at E is the vapor pressure of the liquid in B at the temperature of the bath. Readings at different temperatures can be obtained by merely changing the temperature of bath A and repeating the operation.

A simple dynamic method is shown in Fig. 7. The liquid in question, B, is boiled, after deaeration, under a measured external pressure, and the temperature of the condensing vapor is read from thermometer T. The barometric pressure minus the pressure P is the pressure at which the liquid boils, and this is the vapor pressure of the liquid at the temperature T. By changing the pressure P, the liquid may be boiled at different temperatures, and the vapor pressures at these temperatures thus obtained. The function of bottle C is to condense any escaping vapor and thus prevent the distillation of liquid into the mercury manometer.

A more elaborate dynamic method proposed by Walker involves the saturation of some inert gas, such as nitrogen, with vapor by bubbling a measured quantity of the gas through the liquid at constant temperature, and subsequently condensing out, or absorbing, and weighing the

vapor of the liquid in question. If P_T is the total pressure in the apparatus at saturation, n_g the moles of gas passed through, and $n_v = W_v/M_v$
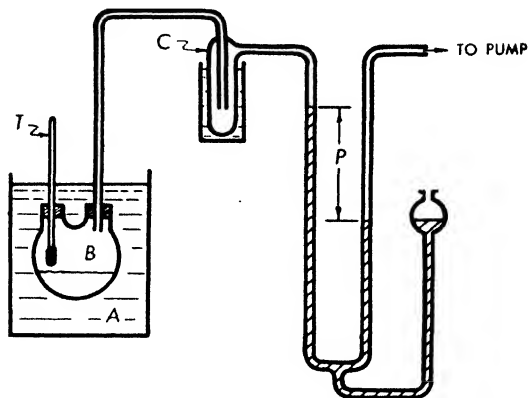
Fig. 7. Boiling Point Method for Determination of Vapor Pressure

the number of moles of vapor collected, then the partial pressure of the vapor P, which is the same as the vapor pressure of the liquid at saturation, is

$$P = \left(\frac{n_v}{n_g + n_v}\right) P_T \tag{19}$$

When the volume of vapor in the inert gas is very small compared to the volume of gas employed, the total volume of gas is essentially the volume of the inert gas V_g, and the whole calculation simplifies to

$$PV_g = n_v RT = \frac{W_v}{M_v} RT \tag{20}$$

where P is the vapor pressure, V_g the volume of gas at temperature T, W_v the weight of vapor collected, and M_v the molecular weight of vapor. This method is as a rule much more tedious than the others mentioned, but with care can be made to yield excellent results. It is especially useful in determinations of partial vapor pressures of mixtures of liquids.

Variation of Vapor Pressure with Temperature. The vapor pressure of a liquid, though constant at a given temperature, increases continuously with increase in temperature up to the critical point of the liquid. Above the critical temperature the liquid no longer exists, and consequently the concept of a saturated vapor pressure is no longer valid. In terms of kinetic theory the increase in vapor pressure with temperature is easily understandable. As the temperature increases, a greater proportion of the molecules acquire sufficient energy to escape from the liquid, and consequently a higher pressure is necessary to

establish equilibrium between vapor and liquid. Above the critical temperature the escaping tendency of the molecules is so high that no applied pressure is sufficient to keep any of them in the liquid state, and the whole mass persists as a gas.

TABLE 2

VAPOR PRESSURES OF LIQUIDS AT VARIOUS TEMPERATURES

(mm Hg)

Temperature °C	CCl₄	C₂H₅OH	CH₃COOH	C₆H₆	H₂O
0	32.9	12.7	3.5	25.3	4.6
10	56.0	24.2	6.4	45.2	9.2
20	91.0	44.5	11.8	75.6	17.5
30	142.3	78.5	20.1	120.2	31.8
40	214.8	133.7	34.2	183.6	55.3
50	314.4	219.9	56.3	271.4	92.5
60	447.4	350.2	88.3	390.1	149.4
70	621.1	541.1	137.9	547.4	233.7
80	843.3	812.9	202.3	753.6	355.1
90	1122.0	1187.0	292.7	1016.1	525.8
100	1463.0	1693.0	417.	1344.3	760.0

The manner in which the vapor pressure varies with temperature is shown in Table 2, and graphically in Fig. 8. The logarithmic nature of the plot is quite characteristic and is exhibited by all liquids. The vapor pressure increases slowly at the lower temperatures, and then quite rapidly, as is shown by the steep rise in the curves. This variation of vapor pressure with temperature can be correlated through the Clausius-Clapeyron equation. We have seen in the last chapter, equation (43), that the variation in pressure with absolute temperature of a liquid-gas system is given by the Clapeyron equation

$$\frac{dP}{dT} = \frac{\Delta H_v}{T(V_g - V_l)} \quad (21)$$

where P is the pressure at absolute temperature T, ΔH_v the heat of vaporization, while V_g and V_l are
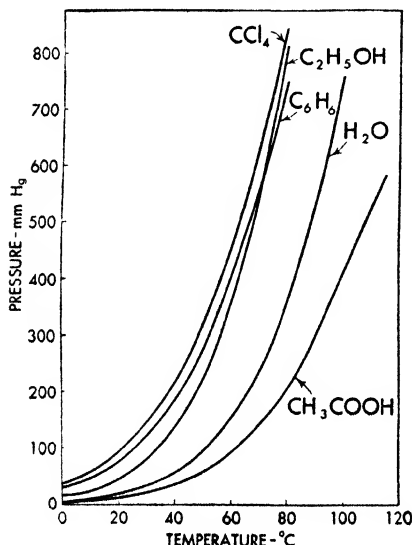
Fig. 8. Variation of Vapor Pressure with Temperature

the volumes of a definite weight of liquid and vapor respectively. At temperatures considerably removed from the critical, V_l is quite small compared to V_g, and may be neglected. Thus, at 100° C, V_g for water is 1671 cc per gram, while V_l is only 1.04 cc per gram. Further, if we assume that the vapor behaves essentially as an ideal gas, then V_g per mole is given by $V_g = RT/P$ and equation (21) becomes

$$\frac{dP}{dT} = \frac{\Delta H_v}{TV_g} = \frac{\Delta H_v P}{RT^2} \tag{22}$$

$$\frac{d \ln P}{dT} = \frac{1}{dT}\left(\frac{dP}{P}\right) = \frac{\Delta H_v}{RT^2} \tag{23}$$

Before this equation can be integrated, ΔH_v must be known as a function of temperature. If we assume as an approximation, however, that over the interval in question ΔH_v remains essentially constant, integration yields

$$\ln P = \frac{\Delta H_v}{R} \int \frac{dT}{T^2} + C'$$

$$= - \frac{\Delta H_v}{R}\left(\frac{1}{T}\right) + C'$$

and
$$\log_{10} P = - \frac{\Delta H_v}{2.303\ R}\left(\frac{1}{T}\right) + C \tag{24}$$

where C' and C are constants of integration.

Equation (24) is one form of the celebrated *Clausius-Clapeyron equation*. It predicts that the logarithm of the vapor pressure should be a function of the reciprocal of the absolute temperature. Further, comparison of the equation with the equation of a straight line, namely, $y = mx + b$, suggests that if $\log_{10} P$ for any liquid is plotted against $1/T$, the plot should be a straight line with slope $m = (- \Delta H_v/2.303\ R)$, and $y -$ intercept $b = C$. That this is in accord with the facts may be seen from Fig. 9, where the same data as were plotted in Fig. 8 are now plotted with $\log_{10} P$ vs. $1/T$. From the slopes of the lines the heats of vaporization of the various liquids may be calculated, for

$$\text{slope} = m = \frac{- \Delta H_v}{2.303\ R} \tag{25}$$

and consequently,

$$\Delta H_v = - 2.303\ Rm = - 4.57\ m \text{ cal/mole} \tag{26}$$

The heat of vaporization is obtained in calories per mole when the value of R used is in calories per mole per degree, namely, $R = 1.987$.

It will be observed that all the slopes in Fig. 9 are negative, and hence all the heats of vaporization will be positive. The heat of vaporization of a liquid thus calculated will not be the heat corresponding to any one temperature, but it will be, rather, the mean value of the heat of vaporization over the temperature interval considered.

To obtain C in equation (24) it is best to substitute in the equation the calculated value of ΔH_v and a value of $\log_{10} P$ and $1/T$, corresponding to a point on the line, and solve for C. Once ΔH_v and C for a given liquid are known, the vapor pressure of the liquid at any temperature over the range of the equation can easily be calculated by merely substituting the desired value of T.

A word about the units of the various terms in equation (24) is in order.

Fig. 9. Plot of $\log_{10} P$ vs. $1/T$ for Several Liquids

Since ΔH_v and R are both in calories, the first term on the right-hand side of the equation is independent of the units in which P is expressed. C, however, is not, and its magnitude will depend on the units of P. Consequently, in setting up an equation for the vapor pressure, it is essential to state clearly whether P is expressed in atmospheres, mm of mercury, or some other unit. Similarly, in using an equation from some reference source, attention should be paid to the units in which the equation is expressed.

An alternate form equation (24) may be obtained by integrating equation (23) between the limits P_1 and P_2 corresponding to the temperatures T_1 and T_2. Then,

$$\int_{P_1}^{P_2} d \ln P = \frac{\Delta H_v}{R} \int_{T_1}^{T_2} \frac{dT}{T^2}$$

$$\ln \frac{P_2}{P_1} = \frac{\Delta H_v}{R} \left[-\frac{1}{T} \right]_{T_1}^{T_2}$$

$$= \frac{\Delta H_v}{R} \left[\frac{T_2 - T_1}{T_1 T_2} \right]$$

or
$$\log_{10} \frac{P_2}{P_1} = \frac{\Delta H_v}{2.303\, R} \left[\frac{T_2 - T_1}{T_1 T_2} \right] \tag{27}$$

Equation (27) permits the calculation of ΔH_v from the values of the vapor pressure at two temperatures; or, when ΔH_v is known, P at some desired temperature may be calculated from a single available vapor pressure at a given temperature.

Example: At 373.6° K and 372.6° K the vapor pressures of H_2O are 1.018 and 0.982 atmospheres respectively. What is the heat of vaporization of water? Employing equation (27), we have

$$\log_{10} \frac{P_2}{P_1} = \frac{\Delta H_v}{2.303\, R} \left[\frac{T_2 - T_1}{T_1 T_2} \right]$$

$$\log \frac{1.018}{0.982} = \frac{\Delta H_v}{2.303 \times 1.987} \left[\frac{373.6 - 372.6}{373.6 \times 372.6} \right]$$

$$\Delta H_v = 9790 \text{ cal/mole}$$
$$= 540 \text{ cal/gram}$$

The experimentally observed value for the heat of vaporization of water at 373.15° K is 538.7 calories per gram.

Equation (24) will not be strictly valid over wide temperature ranges, particularly because of the assumption of constancy of ΔH_v. When data are available on the variation of ΔH_v with temperature, equation (23) may be integrated to give better agreement with experiment and over a much wider temperature range. The manner in which such integrations are carried out will be explained in detail in the chapter on Thermochemistry. When such data are not available, vapor pressure-temperature data are usually correlated by empirical equations. Of these the Kirchhoff equation is especially useful, being valid over a considerable temperature range and even up to the critical point. The equation is

$$\log P = A - \frac{B}{T} + C \log T \tag{28}$$

where A, B, and C are empirically fitted constants. .

The Heat of Vaporization of Liquids. The heats of vaporization of liquids may also be measured directly in a calorimeter by condensing a definite weight of vapor and observing the temperature rise of the calorimeter or by supplying to a liquid a definite amount of electrical energy and measuring the weight of liquid vaporized thereby. The heats of vaporization of several liquids thus obtained at various temperatures are shown in Table 3. The heat of vaporization of a liquid decreases in general with increase in temperature and becomes zero at the critical temperature. For acetic acid, however, ΔH_v increases at first, goes through a maximum at about 120° C, and then decreases with increase

in temperature. In magnitude the heats of vaporization of various liquids differ widely. Especially is the heat of vaporization of water abnormally high, a fact which speaks for the complexity of the liquid.

TABLE 3

HEATS OF VAPORIZATION OF LIQUIDS [1]

(Cal/gram)

Temperature (° C)	H_2O	CH_3OH	CH_3COOH	C_2H_5OH	Ethyl Ether	CCl_4
0	595	285	—	220	93	52
20	585	280	87.3	218	88	—
40	574	273	—	215	83	—
60	563	265	93.3	210	78	—
80	552	253	—	203	72	46
100	539	241	95.6	194	67	44
120	526	227	—	182	61	42
140	512	213	94.5	170	54	40
160	496	194	—	156	44	38
180	478	175	88.5	136	30	35

The Boiling Point of a Liquid. The *normal* boiling point of a liquid is the temperature at which the vapor pressure of the liquid equals 760 mm Hg pressure, or 1 atmosphere. However, a liquid can be made to boil at any temperature between its freezing point and the critical temperature by merely raising or lowering, as the case may be, the external pressure on the liquid. Therefore, it may be stated in general that the boiling point of a liquid is the temperature at which the vapor pressure of a liquid becomes equal to the external pressure acting upon the surface of the liquid. Boiling is characterized by the formation within the liquid of bubbles of vapor which rise and escape into the vapor phase above the surface of the liquid.

The change in boiling point produced by a certain change in pressure may be estimated with the aid of the Clausius-Clapeyron equation. If ΔH_v for the liquid is known, and if T_1 is the boiling point at pressure P_1, the boiling point T_2 at pressure P_2 follows from equation (27). Occasionally it is necessary to correct a boiling point observed at a pressure other than 760 mm Hg to the normal boiling point, or 760 mm Hg. This may be done more simply by taking the reciprocal of equation (22) and considering instead of dT and dP the finite temperature and pressure intervals ΔT and ΔP. Then

$$\frac{\Delta T}{\Delta P} = \frac{RT^2}{\Delta H_v P} \tag{29}$$

[1] International Critical Tables, McGraw-Hill Book Company, Inc., New York, 1929, Vol. V.

ΔT is the change in temperature produced by the pressure change ΔP. Since, as we shall see below, $(\Delta H_v)/T$ at the normal boiling point is a constant and P is not far from 760 mm Hg, $(RT)/(\Delta H_v P) = C$ may be considered to be a constant near the boiling point, and hence:

$$\frac{\Delta T}{\Delta P} = \left(\frac{RT}{\Delta H_v P}\right) T = CT$$
$$\Delta T = CT\Delta P = CT(760 - P) \tag{30}$$

P is the pressure at which the liquid has the observed boiling point T. Equation (30) gives the correction which must be *added* to T in order to obtain the normal boiling point, T_b. When $P < 760$ mm, ΔT is positive and $T_b > T$; when, however, $P > 760$ mm, ΔT is negative and $T_b < T$.

The constant C may be calculated for any liquid from the relation

$$C = \frac{RT_b}{\Delta H_v P} = \frac{RT_b}{\Delta H_v(760)} \tag{31}$$

For many liquids C may be taken as 0.00012. We obtain then the frequently employed Sydney Young equation,

$$\Delta T = 0.00012\, T(760 - P) \tag{32}$$

Thus, for a liquid boiling at 55.4° C under a pressure of 740 mm, the boiling point correction would amount to 0.8° by equation (32), and hence the normal boiling point would be 56.2° C.

Trouton's Rule and Modifications. At the normal boiling point of a liquid P is 1 atmosphere, $\ln P = 0$, and the integrated Clausius-Clapeyron equation (24) becomes

$$\frac{\Delta H_v}{RT_b} = C'$$
$$\frac{\Delta H_v}{T_b} = RC' = \text{constant} \tag{33}$$

Trouton's rule states that the ratio of the molar heat of vaporization of a liquid to its normal boiling point on the absolute scale is a constant the same for all liquids. The value of the constant is usually taken as about 21. That this rule is only approximate may be seen from Table 4. It holds fairly well for the so-called normal liquids, such as methane, hexane, and benzene, but is considerably in error for the liquids that are considered to be associated, such as water, the alcohols, and the acids.

Hildebrand in 1915 suggested that better results are to be expected if all liquids were compared at equal densities rather than at equal pressures. Actually somewhat better concordance is observed under these

TABLE 4

TROUTON'S CONSTANTS FOR LIQUIDS [1]

Liquid	ΔH_v per gram	Normal B.P. (° C)	Trouton Constant
Hexane	79.5	69	20.0
Oxygen	51	− 183	18.2
Benzene	94.6	79.6	20.8
Carbon tetrachloride	46.4	76.8	20.4
Ether	84.1	34.6	20.2
Acetone	124.3	56.5	21.9
Water	539	100.0	26.0
Ethyl alcohol	205	78.4	26.7
Formic acid	120	100.8	14.8
Acetic acid	97.2	118.1	14.8
Propyl alcohol	164	97.8	26.6
Glycol	191	197.4	25.1

conditions, but associated liquids still behave anomalously. Two additional equations of the same type may be mentioned, one due to Nernst,

$$\frac{\Delta H}{T_b} = 9.5 \log_{10} T_b - 0.007 \, T_b \tag{34}$$

and another due to Bingham,

$$\frac{\Delta H}{T_b} = 17 + 0.011 \, T_b \tag{35}$$

Both suffer from the same shortcomings as the simpler Trouton's rule.

Surface Tension of Liquids. Within the body of the liquid a molecule is acted upon by molecular attractions which are distributed more or less symmetrically about the molecule. At the surface, however, a molecule is only partially surrounded by other molecules, and as a consequence it experiences only an attraction toward the body of the liquid. This latter attraction tends to draw the surface molecules inward, and in doing so makes the liquid behave as if it were surrounded by an invisible membrane. This behavior of the surface, called *surface tension*, is the effect responsible for the resistance a liquid exhibits to surface penetration, the nearly spherical shape of falling water droplets, the spherical shape of mercury particles on a flat surface, the rise of liquids in capillary tubes, and the flotation of metal foils on liquid surfaces. From a purely mechanical point of view surface tension may be thought of as due to the tendency of a liquid to reduce its surface to a point of minimum

potential surface energy, a condition requisite for stable surface equilibrium. Since a sphere has the smallest area for a given volume, the tendency of a liquid particle should be to draw itself into a sphere due to the action of surface tension, as is actually the case.

Since the natural tendency of a liquid is to decrease its surface, any increase in surface can only be accomplished with the expenditure of work. Consider a liquid film contained within
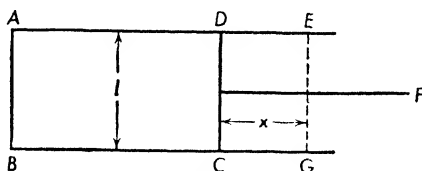
Fig. 10

a rectangular wire frame, $ABCD$, as shown in Fig. 10. The side CD is movable. If a force F is required to move the wire CD against the force of surface tension acting in the film along CD, the work done, w, in moving the wire from CD to EG is

$$w = Fx \qquad (36)$$

The force acting, F, must, of course, be balanced by the force of surface tension along CD. If we designate by γ the force *per centimeter* along CD, and since there are two surfaces to the film,

$$F = 2\,\gamma l \qquad (37)$$

and

$$w = Fx = 2\,\gamma lx \qquad (38)$$

From equation (37) γ may be defined as $\gamma = F/2\,l$, or the *force in dynes acting along 1 cm length of surface*. However, $2\,lx$ is the area of new surface of liquid generated by CD, ΔA, and hence equation (38) becomes also

$$w = \gamma(2\,lx) = \gamma \Delta A$$

and

$$\gamma = \frac{w}{\Delta A} \qquad (39)$$

Consequently, γ may be considered also as the *work in ergs necessary to generate a square centimeter of surface area*, and is, therefore, referred to frequently as the free surface energy of a liquid per square centimeter of area.

Measurement of Surface Tension. The surface tension is a characteristic property of each liquid and differs greatly in magnitude for different liquids. Of the various methods available for measuring surface tension, such as the tensiometer, drop weight, bubble pressure, or capillary rise methods, the last is by far the most important and is considered the standard. The capillary rise method for the estimation of surface tension is based on the fact that most liquids when brought into contact with a fine glass capillary tube will rise in the tube to a level above that of the liquid outside the tube. This will occur only when the

liquid "wets", i.e., adheres to it. If the liquid does not "wet" glass, as mercury for instance, the level inside the capillary will fall below that outside, and the mercury will exhibit a convex surface, as against the concave surface in the first instance.

To understand the theory of the capillary rise method, consider a fine capillary tube of uniform radius r immersed in a vessel containing a liquid that wets glass (Fig. 11). By wetting the inner wall of the capillary the surface of the liquid is increased. To decrease its free surface the liquid must rise within the capillary. As soon as this happens, however, the glass is again wet, and again the liquid draws itself upward. This process does not continue indefinitely, but stops when the force of surface tension acting upward becomes equal to the force due to the column of liquid acting downward. If we call γ the surface tension in dynes per centimeter of inner circumference, and consider the force to be acting at an angle θ, called the *contact angle*, with the vertical, then the force due to surface tension is

Fig. 11. Capillary Rise Method for Determination of Surface Tension

$$F_1 = 2\pi r\gamma \cos\theta$$

This force is balanced by that due to the column of liquid of height h, or

$$F_2 = \pi r^2 h d g$$

where d is the density of the liquid and g is the acceleration of gravity in centimeters per second per second. Therefore, since at equilibrium $F_2 = F_1$,

$$2\pi r\gamma \cos\theta = \pi r^2 h d g$$

and

$$\gamma = \frac{\pi r^2 h d g}{2\pi r \cos\theta} = \frac{r h d g}{2\cos\theta} \qquad (40)$$

For most liquids which wet glass θ is essentially zero, and $\cos\theta = 1$. Then

$$\gamma = \frac{r h d g}{2} \qquad (41)$$

and γ may be calculated provided we know the radius of the capillary, the density of the liquid, and the height to which the liquid will rise in the capillary.

For precise work two corrections must be applied, one for the volume of the meniscus and another for the density of the gas above the liquid. Under these conditions equation (41) becomes

$$\gamma = \frac{\left(h + \dfrac{r}{3}\right)(d_l - d_v)rg}{2} \tag{42}$$

where d_l is the density of the liquid and d_v the density of the gas above the liquid.

In the tensiometer method a platinum fork or ring is immersed in the liquid to be tested, and the force necessary to separate the fork or ring from the liquid surface is then measured. The force required for this operation can be related to the surface tension of the liquid. In the drop weight method, on the other hand, the liquid whose surface tension is to be determined is allowed to pass very slowly through a calibrated capillary tip so as to form an approximately spherical drop. When the drop has attained a definite weight, depending on the size of the capillary tip and the surface tension of the liquid, a portion of it will detach itself and fall. Harkins and Brown [1] have shown that the surface tension of the liquid may be determined from the weight of the falling drop W, and the radius of the capillary tip by the relation

$$\gamma = \frac{Wg}{2\,\pi rf} \tag{43}$$

where f is a complicated function of $(r/V^{1/3})$, V being the volume of the drop. For the nature of the function f and details of the method the student is referred to Harkins's papers on the subject.

Variation of Surface Tension with Temperature. The surface tension of all liquids decreases as the temperature is increased and becomes zero at the critical temperature. The variation of surface tension with temperature may be represented by the Ramsay-Shields equation, namely,

$$\gamma \left(\frac{M}{d_l}\right)^{2/3} = k(t_c - t - 6) \tag{44a}$$

where γ is the surface tension at temperature t, and M, d_l, and t_c are the molecular weight, density, and critical temperature of the liquid respectively. Since M/d_l is the molar volume of the liquid, $(M/d_l)^{2/3}$ is a quantity proportional to the molar surface of the liquid, and hence the product on the left of equation (44) is frequently referred to as the *molar surface energy*. k is a constant which is supposed to be independent of temperature.

[1] Harkins and Brown, J. Am. Chem. Soc., **41**, 499 (1919).

Equation (44a) has been found to be valid for many liquids up to within 30° to 50° of the critical temperature. On rearranging the equation to

$$\gamma \left(\frac{M}{d_l}\right)^{2/3} = -kt + k(t_c - 6) \tag{44b}$$

we see that a plot of $\gamma(M/d_l)^{2/3}$ vs. t should be a straight line with slope equal to $-k$, and k may thus be evaluated without a knowledge of t_c. For many liquids k has been found to be nearly the same and equal approximately to 2.12. For certain other liquids, however, abnormally low values of k are observed. Ramsay and Shields suggested that those liquids which give values of k in the vicinity of 2.12 may be considered to have the normal molecular weight in the liquid state, while those with abnormally low values of k are *associated* in the liquid state, i.e., have molecular weights higher than is to be expected from their simple formulas. Significantly enough, the liquids that obey Trouton's rule are those which yield normal values of k, while such liquids as water, alcohol, and acetic acid, which do not obey Trouton's rule, show values of k considerably lower than 2.12.

However, in several instances, benzylamine and amyl stearate for example, k values *higher* than 2.12 have been observed, suggesting that these liquids are *dissociated* in the liquid state. From a consideration of the nature of these liquids this conclusion seems highly doubtful. Another shortcoming of the Ramsay-Shields equation is that it predicts that γ will be zero at $t = (t_c - 6)$ and that it will become negative at the critical temperature. To obviate the last difficulty Katayama suggested the modified equation

$$\gamma \left(\frac{M}{d_l - d_v}\right)^{2/3} = k(t_c - t) \tag{45}$$

where d_v is the density of the vapor at temperature t. In this equation $\gamma = 0$ when $t = t_c$.

Various other equations have been proposed to represent the variation of surface tension with temperature. Of these the McLeod (1923) equation,

$$\gamma = C(d_l - d_v)^4 \tag{46}$$

where C is a constant characteristic of each liquid, is of interest because it was employed by Sugden as the starting point in his development of the theory of the parachor. The relation of the parachor to molecular structure will be discussed in Chapter XXIII.

In Table 5 are given the surface tensions of several liquids at different temperatures. The value of k in the Ramsay-Shields formula which best fits these data is indicated for each liquid.

TABLE 5

SURFACE TENSION OF LIQUIDS AT VARIOUS TEMPERATURES [1]

(Dynes/cm)

Liquid	0° C	20° C	40° C	60° C	80° C	100° C	k
Water	75.64	72.75	69.56	66.18	62.61	58.85	1.0–1.2
C_2H_5OH	24.05	22.27	20.60	19.01	--	—	1.0
CH_3OH	24.5	22.6	20.9	—	—	15.7	1.0
CCl_4	—	26.8	24.3	21.9	—	—	2.21
Ethyl ether	—	17.0	—	—	—	7.97	2.25
Acetone	26.2	23.7	21.2	18.6	16.2	—	1.9
Toluene	30.74	28.43	26.13	23.81	21.53	19.39	2.2
Benzene	31.6	28.9	26.3	23.7	21.3	—	2.22

The Viscosity of Liquids. Liquids exhibit much greater resistance to flow than gases, and consequently they have much higher viscosity coefficients. The viscosity coefficients of gases increase with temperature, while just the reverse is generally true of liquids, the viscosity of most liquids decreasing with rising temperature. Again, we have seen that the viscosity coefficients for gases at moderate pressures are essentially independent of pressure, while increase of pressure leads to an increase in viscosity for liquids.

Most methods employed for the measurement of the viscosity of liquids are based on either the Poiseuille or Stokes equations. Poiseuille in 1842, after studying the flow of fluids through capillary tubes, proposed the following equation for the coefficient of viscosity of a fluid:

$$\eta = \frac{\pi P r^4 t}{8 \, lV} \tag{47}$$

Here V is the volume of liquid of viscosity η which flows in time t through a capillary tube of radius r and length l under a pressure head of P dynes per square centimeter. This equation has been verified repeatedly. To determine the viscosity of a liquid by this equation it is not always necessary to measure all the quantities indicated when once the viscosity of some reference liquid, usually water, is known with accuracy. If we measure the time of flow of the same volume of two different liquids through the same capillary, then according to the Poiseuille equation the ratio of the viscosity coefficients of the two liquids is given by

$$\frac{\eta_1}{\eta_2} = \frac{\pi P_1 r^4 t_1}{8 \, lV} \cdot \frac{8 \, lV}{\pi P_2 r^4 t_2} = \frac{P_1 t_1}{P_2 t_2}$$

[1] International Critical Tables, McGraw-Hill Book Company, Inc., New York, 1928, Vol. IV.

Since the pressures P_1 and P_2 are proportional to the densities of the two liquids d_1 and d_2, we may write also

$$\frac{\eta_1}{\eta_2} = \frac{P_1 t_1}{P_2 t_2} = \frac{d_1 t_1}{d_2 t_2} \qquad (48)$$

Consequently, once d_1, d_2, and η_2 are known, determination of t_1 and t_2 permits the calculation of η_1, the viscosity coefficient of the liquid under consideration.
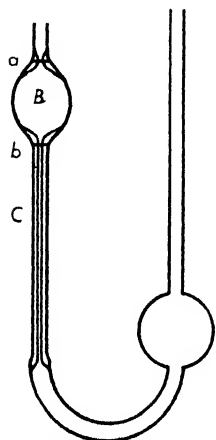
The quantities t_1 and t_2 in equation (48) are most conveniently measured with an Ostwald viscosimeter (Fig. 12). A definite quantity of liquid is introduced into the viscosimeter immersed in a thermostat and is then drawn up by suction into bulb B until the liquid level is above the mark a. The liquid is then allowed to drain, and the time necessary for the liquid level to fall from a to b is measured with a stopwatch. The viscosimeter is now cleaned, the reference liquid added, and the whole operation repeated. In this simple manner t_1 and t_2 are obtained, and the viscosity of the liquid is calculated by equation (48).

Fig. 12. Ostwald Viscosimeter

Stokes's law is concerned with the fall of bodies through fluid media. A *spherical* body of radius r and density d, falling under gravity through a fluid of density d_m, is acted on by the gravitational force f_1,

$$f_1 = \frac{4}{3} \pi r^3 (d - d_m) g \qquad (49)$$

where g is the acceleration of gravity. This force, which tends to accelerate the body falling through the fluid medium, is opposed by frictional forces within the medium caused by the motion of the body which increase with increase in velocity. Eventually a certain maximum rate of fall is reached at which the frictional forces become equal to the gravitational force, and thereafter the body will continue to fall with a *constant velocity* v. Sir George G. Stokes showed that, for a spherical body falling under the conditions of constant maximum velocity, the force of friction, f_2, depends on the viscosity coefficient and is given by

$$f_2 = 6 \pi r \eta v \qquad (50)$$

Equating the gravitational and frictional forces, we see that

$$\frac{4}{3}\pi r^3 (d - d_m)g = 6\pi r\eta v$$

$$\eta = \frac{2\, r^2(d - d_m)g}{9\, v} \tag{51}$$

This equation, known as *Stokes's law*, is applicable to the fall of spherical bodies in all types of fluid media provided the radius of the falling body r is large compared to the distance between the molecules of the fluid. When r is smaller than the distance between molecules there is a tendency for the falling body to "drop" or "channel," and the equation is no longer applicable.

Stokes's law is the basis of the falling sphere viscosimeter shown in Fig. 13. The viscosimeter tube V is filled with the liquid under test, and is immersed in a thermostat B at the desired temperature. A steel ball, of density d and a diameter suitable to give a slow rate of fall, is now dropped through the neck of the tube, and the time of fall between the marks c and d is determined with a stopwatch. The viscosity coefficient of the liquid in V is calculated from the equation [1]

$$\eta = \frac{2\, r^2(d - d_m)g}{9\, \dfrac{s}{t}\left(1 + \dfrac{2.4\, r}{R}\right)} \tag{52}$$

where s is the distance of timed fall (i.e., between c and d), d_m is the density of the liquid, r is the radius of the ball, t is the time required for fall through the dis-

Fig. 13. Falling Sphere Viscosimeter

tance s, and R is the radius of the viscosimeter tube V. The last term in the denominator is a correction for the influence of the size of the container, and becomes negligible when R is very much greater than r; then equation (52) reduces to the simple Stokes's law equation (51). From either equation (52) or (51) the viscosity coefficient of a liquid relative to a standard is given by

$$\frac{\eta_1}{\eta_2} = \frac{(d - d_{m_1})t_2}{(d - d_{m_2})t_1} \tag{53}$$
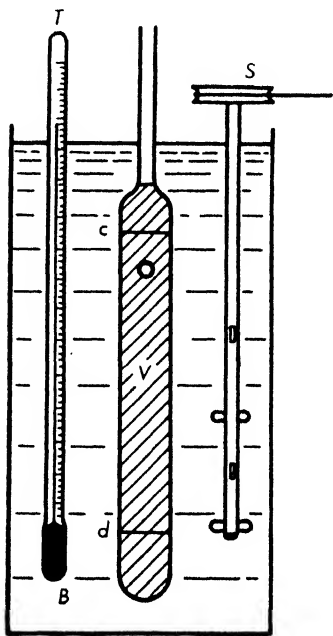
[1] Friend, "Textbook of Physical Chemistry," J. B. Lippincott Company, Philadelphia, 1933, Vol. I.

Hence, the relative viscosity of a liquid may be determined from the times of fall of the same ball and the densities of the two liquids without the knowledge of s, r, or R.

Variation of Viscosity with Temperature. Table 6 gives the viscosity coefficients in centipoises of several liquids at various temperatures. With very rare exceptions (liquid carbon dioxide at low temperatures), the viscosity of a liquid decreases with increase in temperature. Various equations have been proposed to represent η as a function of T, of which possibly the one due to Shepard is simplest, namely,

$$\log \eta = \frac{A}{T} + B \qquad (54)$$

A and B are constants, and T is the absolute temperature. This equation holds quite well for a large number of pure liquids. The similarity of this equation to the Clausius-Clapeyron equation for the vapor pressure of a liquid has led Friend to propose the equation

$$\log \eta = \frac{A}{T} + B \log T + C \qquad (55)$$

which is analogous to the equation proposed by Kirchhoff for the variation of vapor pressure with temperature, and gives very good results. It should be emphasized, however, that, unlike the vapor pressure equations, the viscosity equations have as yet no theoretical justification and must be considered to be purely empirical.

TABLE 6

VISCOSITY COEFFICIENTS OF LIQUIDS [1]

(Centipoises)

Liquid	0° C	20° C	40° C	60° C	80° C	100° C
H_2O	1.794	1.009	0.654	0.470	0.357	0.284
CH_3OH	0.808	0.593	0.449	0.349		
$CHCl_3$	0.699	0.563	0.464	0.389		
C_8H_{18}	0.7060	0.5419	0.4328	0.3551	0.2971	
Toluene	0.7719	0.5903	0.4713	0.3874		
C_2H_5OH	1.772	1.114	1.716	1.647	1.581	
Benzene	0.900	0.647	0.492	0.389		
Nitrobenzene	3.083	2.013	1.438	1.094	0.875	0.705

The viscosity of liquids is a question of considerable industrial importance, especially in the petroleum industry. The use of an oil for lubri-

eating purposes is determined to a large extent by its viscosity and the manner in which the viscosity varies with temperature. In practice various methods of measuring viscosity have been evolved, and various schemes for representing the variations of viscosity with temperature have been proposed. These, however, are beyond the scope of this book.

REFERENCES FOR FURTHER READING

1. G. Barr, "Monograph on Viscometry," Oxford University Press, Oxford, 1931.
2. E. C. Bingham, "Fluidity and Plasticity," McGraw-Hill Book Company, Inc., New York, 1922.
3. J. N. Friend, "Textbook of Physical Chemistry," J. B. Lippincott Company, Philadelphia, 1933, Vol. I.
4. E. Hatschek, "The Viscosity of Liquids," George Bell & Sons, Ltd., London, 1928.
5. K. Jellinek, "Lehrbuch der physikalischen Chemie," Ferdinand Enke, Stuttgart, 1928, Vol. I.
6. Reilly, Rae, and Wheeler, "Physico-Chemical Methods," D. Van Nostrand Company, Inc., New York, 1939.
7. W. Swietoslawski, "Ebulliometry," Jagellonian Univ. Press, Krakow, 1937.
8. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931.
9. Willows and Hatschek, "Surface Tension and Surface Energy," The Blakiston Company, Philadelphia, 1923.

PROBLEMS

1. Calculate the van der Waals constants for C_2H_6 from the critical temperature and pressure listed in Table 1. Using the constants thus calculated find the pressure exerted by 10 g of C_2H_6 when contained in a liter flask at 13° C.
2. Calculate the critical density of methyl alcohol from the following data:

$t°$ C	P (atm)	$d_{liq.}$ (g/cc)	$d_{vap.}$ (g/cc)
150	13.57	0.6495	0.01562
225	61.25	0.4675	0.1003

The critical temperature is 240.0° C.
3. Compare the reduced pressures of N_2 and NH_3 when each exerts a pressure of 100 atm.   *Ans.* N_2: 2.99; NH_3: 0.90
4. Compare the reduced temperatures of ethylene and H_2 at 27° C.
5. In measuring the vapor pressure of a liquid by means of the isoteniscope, the height of the Hg in the manometer was found to be 53.32 cm at 40° C, and 39.40 cm at 55° C. The barometric pressure was 741.0 mm. What are the vapor pressures of the liquid at the two temperatures?
6. In measuring the vapor pressure of ethanol by the gas saturation method, the following data were taken:

Volume of N_2 at 740 mm and 30° C	= 5.6 liters
Barometric pressure	= 740 mm
Temperature	= 30° C
Loss in weight of alcohol	= 1.193 g

Find the vapor pressure of ethanol at 30° C.   *Ans.* 78.2 mm Hg

7. The vapor pressure of ethanol is 135.3 mm at 40° C and 542.5 mm at 70° C. Calculate the molar heat of vaporization and the vapor pressure of ethanol at 50° C. Compare your results with those given in the literature.

Ans. 9880 cal/mole

8. The heat of vaporization of ethyl ether is 83.9 cal/gram at its boiling point, while its vapor pressure at 30° is 647.3 mm. What will be the vapor pressure at 0° C?

9. At the normal boiling point of isopropanol, 82.3° C, its heat of vaporization is 159 cal/gram. Calculate the vapor pressure of this liquid at 27° C.

10. For $CHCl_3$ the normal boiling point is 61.5° C, while the heat of vaporization is 59.0 cal/gram. At what temperature will $CHCl_3$ boil under a pressure of 700 mm?

11. A liquid is observed to boil at 120° C under a pressure of 725 mm. Its molar heat of vaporization is 8200 cal/mole. Calculate the normal boiling point of the liquid.

12. The normal boiling point of C_6H_5Br is 156.15° C. Using Trouton's rule, find the vapor pressure at 100° C, and compare with the observed value of 141.1 mm.

13. The vapor pressure of CH_3Cl between $-47°$ C and $-10°$ C can be represented by the equation:

$$\log_{10} P_{mm} = \frac{-1149}{T} + 7.481$$

What is the heat of vaporization of this liquid in calories per gram?

14. In measuring the heat of vaporization of a liquid, the following data were obtained:

$$\begin{aligned}
\text{Wt. of } H_2O \text{ in calorimeter} &= 450 \text{ g} \\
\text{Initial temperature of } H_2O &= 25.012° \text{ C} \\
\text{Final temperature of } H_2O &= 26.525° \text{ C} \\
\text{Weight of condensed vapor} &= 13.52 \text{ g}
\end{aligned}$$

Find the heat of vaporization in calories per gram.   *Ans.* 50.4 cal/g

15. The radius of a given capillary is 0.105 mm. A liquid whose density is 0.80 g/cubic centimeter rises in this capillary to a height of 6.25 cm. Calculate the surface tension of the liquid.

16. From the following data calculate the critical temperature of CO_2:

Temperature ° C	Density (g/cc)	Surface Tension (dynes/cm)
0	0.927	4.5
20	0.772	1.16

17. In measuring the surface tension of a liquid by the drop-weight method, 12 drops of the liquid falling from a tip whose diameter is 0.8 cm are found to weigh 0.971 g. If $f(r/V^{1/3}) = 0.6$ under these conditions, what is the surface tension in dynes per centimeter?   *Ans.* 26.29 dynes/cm

18. A steel ball of density 7.9 g/cubic centimeter and 4 mm diameter requires 55 sec to fall a distance of 1 meter through an oil of density 1.1 g/cubic centimeter. Neglecting the correction due to the radius of the viscosimeter tube, calculate the viscosity of the oil in poises.

19. The time of efflux of H_2O through an Ostwald viscosimeter is 1.52 min. For the same volume of an organic liquid of density 0.800 g/cubic centimeter the time is 2.25 min. Find the viscosity of the liquid relative to that of water, and its absolute viscosity in millipoises. The temperature is 20° C.

CHAPTER IV

The Solid State

Introduction. Solids differ from liquids and gases in possessing both definite volume and definite shape. Whereas a liquid always adopts the shape of its container, a solid has sufficient rigidity to enable it to withstand much more successfully the forces of shearing, compression, tension, and torsion. The geometric stability of a solid is not due to any difference in compactness between the solid and liquid states, for the density of a substance in the solid state may actually be less than that of the corresponding liquid, as in the case of ice and water, for instance. The definite shape of a solid is to be ascribed rather to the fact that the structural units, instead of being in random motion like the molecules of a liquid or gas, are confined to a definite position of equilibrium within the crystal of the solid, a position about which the particles may vibrate but which they cannot readily leave.

Solid substances are frequently classed as either *crystalline* or *amorphous*. A crystalline solid is one in which the constituent structural units are arranged in a definite geometrical configuration characteristic of the substance. Amorphous substances, on the other hand, although possessing many of the attributes of a solid, such as definite shape, a certain rigidity, and hardness, do not show under test a definite configurational arrangement. For that reason they are not considered to be true solids but rather highly supercooled liquids of very high viscosity. Further, crystalline substances such as ice, sodium chloride, or naphthalene melt sharply at a constant and definite temperature, while amorphous substances like glass or asphalt melt gradually and over a temperature interval. However, under certain conditions an amorphous substance may acquire crystalline characteristics. Thus glass may crystallize on long standing or heating. Again, rubber when stretched exhibits a definite pattern on examination with x rays, an indication of the production of a definite configurational arrangement.

Crystallization and Fusion. A pure liquid on being cooled suffers a decrease in the average translational energy of its molecules, and hence its temperature drops until the freezing point is reached. At this temperature the attractive forces of the molecules are sufficient to overcome

109

the translational energy, and the molecules are forced to arrange themselves in a geometric pattern which is characteristic for each substance. When crystallization starts, heat is evolved in amount equal to the difference in heat content between solid and liquid, namely,

$$\Delta H_c = H_s - H_l \tag{1}$$

where ΔH_c is the *heat of crystallization*, while H_s and H_l are the heat contents of solid and liquid respectively, all per mole. This heat evolution arrests further temperature drop, and the temperature of the mixture of solid and liquid remains constant as long as both phases are present. Further removal of heat results merely in the crystallization of more liquid, until finally the whole mass solidifies; only then does the temperature begin to fall again on cooling.

The freezing point, or the temperature at which solid and liquid are in equilibrium, is constant and definite for each substance. Studies have shown that crystallization is usually initiated by the formation of crystal nuclei within the liquid around which the crystals grow and develop. Most liquids have a tendency to supercool, i.e., to be cooled below the freezing point, before crystal formation sets in. This is especially true when the cooling is rapid. To prevent supercooling and to facilitate the formation of nuclei for crystallization, it is usually efficacious to cool the liquid slowly, stir vigorously, and "seed" the liquid. Seeding involves the addition of a small quantity of the crystals to be formed to act as nuclei. Once crystallization starts, the mass will return to its equilibrium temperature and stay there until crystallization is complete.

The reverse of crystallization is the fusion or melting of the solid. As the pure solid is heated its average vibrational energy increases, until at the melting point some particles are vibrating with sufficient energy to overcome the confining forces. The solid then begins to fuse. The temperature at which this occurs is the *same* as the crystallization temperature. To accomplish further fusion, heat must be supplied to compensate for the loss of the particles with high energy. The amount of heat which must be *absorbed* to accomplish the transition of 1 mole of solid to 1 mole of liquid is known as the *heat of fusion*, ΔH_f. This amount of heat must represent the difference in heat contents of 1 mole of liquid and solid respectively, and must therefore be *equal in magnitude*, but be *opposite in sign*, to the heat of crystallization of the substance, namely,

$$\Delta H_f = (H_l - H_s) = -\Delta H_c \tag{2}$$

Strictly speaking the melting point of a solid is constant only at a fixed confining pressure, and will be displaced by a change in external pressure in a direction which is determined by the relative densities of

solid and liquid. The variation of melting point with pressure is given by the Clapeyron equation,

$$\frac{dT}{dP} = T\,\frac{(V_l - V_s)}{\Delta H_f} \tag{3}$$

where dT is the change in melting point produced by the change in pressure dP, T is the melting point, and V_l and V_s are the molar volumes of liquid and solid respectively. If the volumes are expressed in cubic centimeters and the pressures in atmospheres, ΔH_f must be expressed in cubic centimeter-atmospheres. To convert ΔH_f in calories to this unit it is merely necessary to multiply ΔH_f by 41.2. From equation (3) it is seen that, since ΔH_f and T are always positive, an increase in pressure (dP positive) will result in a rise in melting point if V_l is greater than V_s, and a lowering in melting point if V_s is greater than V_l. In other words, increase in external pressure will favor the state having the higher density.

Equation (3) cannot be integrated unless both ΔH_f and $(V_l - V_s)$ are known as functions of the temperature or pressure. Since such data are usually not available, it is customary to employ the equation in differential form. Its use can be illustrated by an example. The melting point of acetic acid is 16.60° C at atmospheric pressure, while ΔH_f and $(V_l - V_s)$ are 46.42 calories and 0.1601 cc per gram respectively. Designating by M the molecular weight of acetic acid, we obtain for the change in melting point per atmosphere change in pressure:

$$\frac{\Delta T}{\Delta P} = T\,\frac{(V_l - V_s)}{\Delta H_f} = \frac{289.72(0.1601)M}{(46.42 \times 41.2)M}$$
$$= 0.0242 \text{ degree/atmosphere}$$

At 11 atmospheres total pressure, for instance, $\Delta P = 10$, and the melting point at that pressure would be

$$t = 16.60 + 0.0242 \times 10$$
$$= 16.84° \text{ C}$$

It will be observed that the effect of pressure on the melting point is not very large, and for small pressure variations may be disregarded.

Water is the interesting example of a substance having V_s greater than V_l, and hence the melting point of water is lowered by the application of pressure.

Sublimation Pressure of Solids. Many solids when placed in an evacuated space are found to exhibit discernible and measurable vapor pressures. At any given temperature, the vapor pressure established is constant and unique for each substance. The equilibrium vapor pressure

of a solid, known as the *sublimation pressure*, is entirely analogous to the saturated vapor pressure of a liquid. As in liquids, the sublimation pressure represents an equilibrium between the rate of evaporation of atoms or molecules from the surface of the solid and the rate of return of molecules or atoms to the solid.

The effects of the direct evaporation or sublimation of solids are frequently directly observable. In winter snow will slowly disappear even though the temperature remains considerably below freezing. Naphthalene (moth balls) exposed in a room will vanish with time, and the presence of the vapor can be detected by the odor. Iodine is customarily purified by sublimation of the solid and condensation of the vapors. The fact that many solid substances like vanillin or camphor give off characteristic odors speaks for the evaporation of the solids. The presence of such vapors is detectable by odor or is directly measurable. Many other solids, however, exhibit no odor or measurable vapor pressures. Still, from a theoretical point of view it is convenient to ascribe to a solid a sublimation pressure at each temperature, no matter how small it may be, and whether it is measurable or not.

The process of sublimation, like that of vaporization of a liquid, is associated with a thermal change. If our unit is again the mole, then the molar *heat of sublimation* is the number of calories required to vaporize 1 mole of solid at any given temperature. More precisely, the heat of sublimation, ΔH_s, constitutes the difference in the heat contents of vapor and solid, namely,

$$\Delta H_s = H_v - H_s \tag{4}$$

where H_s is the heat content of the solid, while H_v is the heat content of the vapor.

Like the vapor pressure of a liquid, the sublimation pressure of a solid increases with temperature. The effect of temperature on the sublimation pressure is readily deducible from the Clapeyron equation, which for the equilibrium solid-vapor takes the form

$$\frac{dP}{dT} = \frac{\Delta H_s}{T(V_g - V_s)} \tag{5}$$

ΔH_s is the molar heat of sublimation, V_g and V_s the volumes of 1 mole of vapor and solid respectively, and P the sublimation pressure at any absolute temperature T. Since V_s is negligibly small compared to V_g, and if we assume the ideal gas laws to be valid for the vapor, then $V_g = RT/P$, and

$$\frac{dP}{dT} = \frac{\Delta H_s}{T(V_g)} = \frac{\Delta H_s P}{RT^2}$$

$$\frac{d\ln P}{dT} = \frac{\Delta H_s}{RT^2} \tag{6}$$

Integrating this equation on the assumption that ΔH_s is constant, once without limits, and again between the limits P_1 and P_2 at the temperature T_1 and T_2, we obtain

$$\log_{10} P = -\frac{\Delta H_s}{2.303\ R}\left(\frac{1}{T}\right) + C \tag{7a}$$

$$\log_{10}\left(\frac{P_2}{P_1}\right) = \frac{\Delta H_s}{2.303\ R}\left[\frac{T_2 - T_1}{T_1 T_2}\right] \tag{7b}$$

These equations represent integrated forms of the Clausius-Clapeyron equation as applied to sublimation. They are exactly analogous to the equations developed in the preceding chapter for the vapor pressure of a liquid, and are manipulated in the same way. They may be employed to represent the sublimation pressure of a solid as a function of temperature, to evaluate the heat of sublimation when the vapor pressures are known at at least two temperatures, or to calculate the sublimation pressure at a desired temperature when ΔH_s and P at some given temperature are known.

The Heat Capacity of Solids. As early as 1819, Dulong and Petit found that for most of the solid *elements* at room temperature the heat
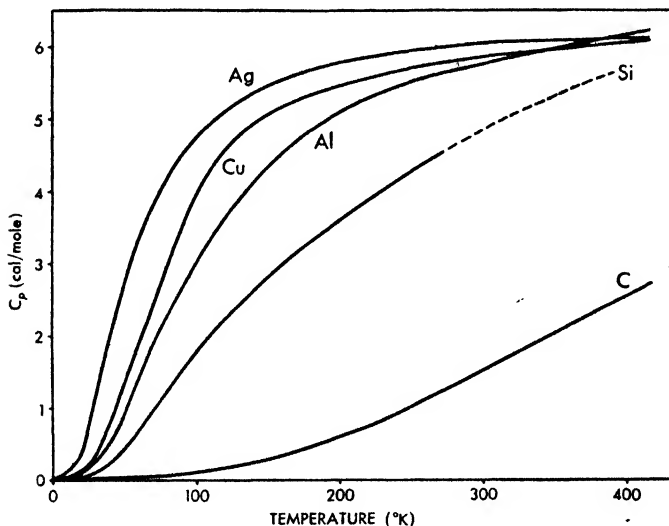
Fig. 1. Variation of Heat Capacity with Temperature

capacity per gram atom was essentially the same and equal to approximately 6.3 calories. Although Dulong and Petit's rule holds approximately for many solid elements, there are instances where the rule breaks down completely. Thus the mean atomic heat capacity of boron is 3.31, beryllium 3.82, silicon 4.77, while carbon (diamond) is only 1.36.

TABLE 1

C_p PER MOLE OF ELEMENTS AT VARIOUS TEMPERATURES
(calories)

Temp. (° K)	Al	C	Cu	Pb	Si	Ag	Na
20	0.07	0.00	0.11	2.62	0.03	0.39	—
40	0.53	0.00	0.81	4.82	0.25	2.00	—
60	1.43	0.018	1.90	5.43	0.68	3.37	4.01
80	2.30	0.048	2.98	5.73	1.25	4.22	4.77
100	3.01	0.084	3.94	5.86	1.78	4.76	5.27
120	3.64	0.144	4.58	5.91	—	5.12	5.61
140	4.12	0.222	4.93	5.96	2.60	5.37	5.82
160	4.53	0.330	5.15	6.01	—	5.55	5.93
180	4.84	0.456	5.30	6.05	—	5.67	6.05
200	5.10	0.594	5.43	6.10	3.62	5.77	6.14
240	5.45	0.924	5.65	6.20	4.14	5.92	6.35
260	5.58	1.116	5.73	6.25	—	5.95	6.39
298	5.74	1.45	5.83	6.31	—	6.02	6.66
323	5.84	1.83	5.88	6.34	—	6.04	6.85
373	6.04	2.24	5.98	6.51	—	6.08	7.40
473	6.34	3.24	6.14	6.88	—	6.14	—

Theoretical considerations suggest that better concordance with Dulong and Petit's rule is to be expected for C_v than C_p; and, G. N. Lewis in 1907 actually showed that C_v for many solid elements whose atomic weight is higher than potassium is equal to 5.9 ± 0.09 calories. This value is in excellent agreement with that predicted in 1871 by Boltzmann on the basis of classical mechanics, namely, $C_v = 3 R = 5.97$ calories. However, even with these refinements Dulong and Petit's rule is still invalid for elements of low atomic weight at room temperature, and for all elements at low temperatures.

The experimental fact is that the heat capacities of solid elements are not constant, as suggested by Dulong and Petit, but decrease gradually with decrease in temperature. The drop is especially pronounced at low temperatures. The variation of the heat capacity with temperature for several elements is given in Table 1, and is shown graphically

in Fig. 1. It will be observed that the heat capacities of elements like aluminum, copper, and silver start at low values and increase very rapidly with temperature, approaching in general the Dulong and Petit value of 3 R at or near room temperature. Carbon and silicon, on the other hand, show a much more gradual increase in heat capacity with temperature, and do not attain the 3 R value until very much higher temperatures. In fact, the heat capacity of carbon does not become 3 R until above 1300° C.

The first satisfactory approach to the theory of specific heats of crystalline monatomic solids was made by Albert Einstein in 1907. Einstein pointed out that the atoms constituting a crystal may be considered to be oscillators executing simple harmonic motion about their mean positions of equilibrium in the crystal. He ascribed to each substance a constant and characteristic frequency of vibration, ν, and postulated that absorption of energy by the oscillators does not take place continuously as called for by classical mechanics, but discontinuously, as postulated in the quantum theory of Planck (see Chapter XXI). With these assumptions Einstein was able to show that the heat capacity per mole at constant volume C_v, at any temperature T, should be given for a monatomic crystalline solid by the equation

$$C_v = 3\,kN \left(\frac{h\nu}{kT}\right)^2 \frac{e^{h\nu/kT}}{(e^{h\nu/kT} - 1)^2} \tag{8}$$

Here N is Avogadro's number, $k = R/N = 1.380 \times 10^{-16}$ ergs/degree, e the base of natural logarithms, h a universal constant known as Planck's constant and equal to 6.624×10^{-27} erg-sec, while ν is the characteristic frequency. This equation predicts that C_v will approach zero at $T = 0$, and that at high temperatures C_v will approach asymptotically the value 3 R. In both these respects the equation is in general agreement with the facts. In the intermediate range, however, the equation gives C_v values considerably lower than those actually observed, although the shape of the curves obtained by plotting C_v against log T according to equation (8) is very similar to the shape of the experimental curves.

The theory of specific heats of solids most successful to date is that of Peter Debye (1912). Debye assumed that a solid does not vibrate with a single characteristic frequency, but may be capable of vibrating with any frequency from zero up to a certain limiting frequency ν_m. By introducing certain principles from the theory of elasticity, and by employing the quantum theory, he was able to derive an equation for C_v as a function of temperature which, although more complicated than Einstein's, is in excellent accord with experiment for a large group of crystalline solids. Like Einstein's theory, Debye's theory predicts that C_v will become zero at $T = 0$, and will approach asymptotically the value 3 R

at very high temperatures. Unlike the Einstein equation, however, the Debye equation reproduces in excellent manner the experimental results in the intermediate temperature ranges not only of elementary crystalline solids but also of such substances as sodium, potassium, lead, and silver chlorides.

An especially valuable contribution of the Debye theory is the fact that it predicts at very low temperatures a linear relation between C_v and T^3, namely,

$$C_v = AT^3 \tag{9}$$

where A is a constant. This equation, known as the *Debye third power law*, has been repeatedly verified. Since experimental determinations of specific heats cannot be carried conveniently below 15 to 20° K, the Debye third power law is employed to estimate the heat capacities of solids between 0 and say 20° K. This extrapolation makes possible the evaluation of entropies according to the third law of thermodynamics, as will be explained in Chapter X.

A serious shortcoming of the Debye theory should be pointed out, however. It accounts only for heat capacities up to $3R$ as an upper limit. Yet certain elements, particularly the alkali metals, reach values of C_v considerably above this limit at high temperatures. The excess absorption of energy is usually ascribed to electrons, of whose displacement by thermal means the simple Debye theory takes no account.

Crystallography. *Crystallography* is the branch of science which deals with the geometry, properties, and structure of crystals and crystalline substances. Geometric crystallography is concerned with the outward spatial arrangement of crystal planes and the geometric shape of crystals, and is based on three fundamental laws, namely,

1. Law of constancy of interfacial angles.
2. Law of rationality of indices.
3. Law of symmetry.

The law of constancy of interfacial angles states that for any given substance the corresponding faces or planes forming the external surface of a crystal always intersect at a definite angle, and that this angle remains constant no matter how the faces develop. Commonly it is observed that the crystal planes are unequally developed so as to produce faces of variable size and shape; but the angle of intersection of any two corresponding faces is always found to be the same for any crystal of the same substance.

For any crystal a set of three coordinate axes can be so chosen that all the faces of the crystal will either intercept these axes at definite distances from the origin, or be parallel to some of the axes, in which case

the intercepts are at infinity. The law of rationality of indices or intercepts, proposed in 1784 by Haüy, states that it is possible to choose along the three coordinate axes unit distances (a, b, c), not necessarily the same length, such that the ratio of the three intercepts of any plane in the crystal is given by $(ma : nb : pc)$, where m, n, and p are either integral whole numbers, including infinity, or fractions of whole numbers. The law may be illustrated with data on crystals of the mineral topaz, $Al_2(FOH)_2SiO_4$, for which four different planes have the parameters,

1. $m = 1$ $n = 1$ $p = 1$
2. $m = 1$ $n = 1$ $p = \infty$
3. $m = 1$ $n = 1$ $p = 2/3$
4. $m = 2$ $n = 1$ $p = \infty$

and hence the intercept ratios are

1. $a : b : c$
2. $a : b : \infty c$
3. $a : b : 2/3\, c$
4. $2\, a : b : \infty c$

For any particular plane these ratios characterize the plane, and may consequently be used to represent it. The coefficients of a, b, and c are known as the *Weiss indices* of a plane. However, Weiss indices are rather awkward in use, and have consequently been replaced by Miller indices, a system of crystal face notation which is practically universally employed now. The Miller indices of a plane are obtained by taking the reciprocals of the Weiss coefficients and multiplying through by the smallest number that will express all the reciprocals as integers. Thus a plane which in the Weiss notation is given by $a : b : \infty c$ becomes in the Miller notation $a : b : 0\, c$, or simply (110), since the order a, b, c is understood. Similarly a face $a : \infty b : 1/4\, c$ becomes (104). As an exercise the student may verify the statement that the four planes mentioned above for topaz are respectively (111), (110), (223), and (120) in the Miller system of crystal face notation.

The third law of crystallography states simply that all crystals of the same substance possess the same elements of symmetry. There are three possible types of symmetry. First, if a crystal can be divided by an imaginary plane passed through its center into two equal portions each of which is a mirror image of the other, the crystal is said to possess a *plane of symmetry*. Second, a crystal is said to possess *line symmetry* if it is possible to draw an imaginary line through the center of the crystal and then revolve the crystal about this line through 360° in such a way as to cause the crystal to appear unchanged two, three, four, or six times. Depending on the number of times the crystal appears unchanged on

revolution, the crystal is said to possess two-, three-, four-, or sixfold symmetry. Finally, a crystal is said to possess a center of symmetry if every face has an identical face at an equal distance on the opposite side of this center. The total number of plane, line, and center symmetries possessed by a crystal is termed the *elements of symmetry* of the crystal.

The Crystal Systems. On the basis of their symmetry crystal forms may be grouped into 32 classes, and these in turn may be referred to six crystal systems. All the crystals belonging to a particular system are characterized by the fact that, although they may not all have the same
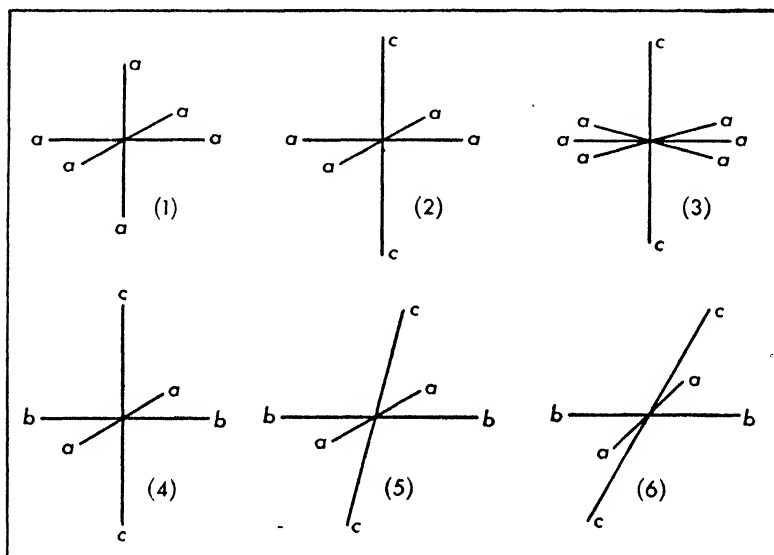
Fig. 2. Axes for Various Crystal Systems

elements of symmetry, they can all be referred to a particular set of crystallographic axes which differ from system to system in length of the various axes and the angles of inclination between axes. Table 2 lists the six crystal systems, their axial characteristics, the maximum symmetry which may be expected in each system, and some examples of substances crystallizing in the various systems. The axes corresponding to these systems are shown in Fig. 2. For a discussion of the various geometric forms which correspond to each of these systems the student must be referred to treatises on the subject.[1]

Such elementary forms as the cube, the octahedron, and the dodecahedron, a figure possessing 12 sides each of which is a rhombus, are all forms corresponding to the regular system, and all possess the maximum

[1] See, for example, Kraus and Hunt, "Mineralogy," second edition, McGraw-Hill Book Company, Inc., New York, 1928, pp. 21–87 and 414–420.

TABLE 2

CRYSTAL SYSTEMS AND THEIR CHARACTERISTICS

System	Axial Characteristics	Symmetry	Substance Typical of System
1. Regular (cubic or isometric)	Three axes at right angles. Unit distances: $a = b = c$	Nine planes Thirteen axes	NaCl KCl Alums Diamond CaF_2 (Fluorspar)
2. Tetragonal	Three axes at right angles, only two of equal length. Unit distances: $a = b \neq c$	Five planes Five axes	TiO_2 (Rutile) $ZrSiO_4$ (Zircon) SnO_2 (Cassiterite)
3. Orthorhombic (rhombic)	Three axes at right angles, but all of different length. Unit distances: $a \neq b \neq c$	Three planes Three axes	KNO_3 Rhombic sulfur K_2SO_4 $BaSO_4$ (Baryte) $PbCO_3$ (Cerrusite)
4. Hexagonal	Three axes of equal length in one plane making angles of 60° with each other, and a fourth axis at right angles to these and of unequal length. Unit distances: $a = b \neq c$	Seven planes Seven axes	PbI_2 Mg Beryl CdS (Greenockite) ZnO (Zincite)
5. Monoclinic	Three axes, all unequal. Two axes at right angles, the third inclined to these at an angle other than 90°. Unit distances: $a \neq b \neq c$	One plane One axis	$Na_2SO_4 \cdot 10\ H_2O$ $Na_2B_4O_7 \cdot 10\ H_2O$ $CaSO_4 \cdot 2\ H_2O$ Monoclinic sulfur
6. Triclinic	Three axes of unequal length, all inclined at angles, and not one 90°. Unit distances: $a \neq b \neq c$	No planes No axes	$CuSO_4 \cdot 5\ H_2O$ $K_2Cr_2O_7$ H_3BO_3

symmetry. A form of lower symmetry, but still in the same system, is the tetrahedron. The cube will be discussed in greater detail further in the chapter.

Properties of Crystals. For gases, liquids, and unstrained amorphous solids, such properties as index of refraction, coefficient of thermal expansion, thermal and electrical conductivity, and rate of solubility are all independent of direction. The same is true of substances crystallizing in the regular system. Such substances exhibiting the same proper-

ties in all directions are said to be *isotropic*. However, for substances crystallizing in the other crystal systems the properties enumerated above may vary according to the axis along which observation is made. Substances exhibiting directional differences in properties are said to be *anisotropic*. For such substances the coefficients of thermal expansion may not only differ in different directions along the crystal, but may actually be positive in one direction and negative in another, as is the case with silver iodide. Again, unlike isotropic solids, anisotropic solids exhibit more than one index of refraction for the same crystal. From the standpoint of optics, anisotropic substances are subdivided into *uniaxial* and *biaxial* crystals. Uniaxial crystals, embracing the tetragonal and hexagonal systems, possess two indices of refraction, according to the axis along which observation is made. Thus ice, which is hexagonal, has the refraction indices 1.3090 and 1.3104, both for sodium-D line. Biaxial crystals, on the other hand, embracing the orthorhombic, monoclinic, and triclinic systems, have three indices of refraction, as for instance the triclinic potassium dichromate, for which the indices of refraction, depending on the axis of observation, are 1.7202, 1.7380, and 1.8197, all for the sodium-D line.

The anisotropic character of certain crystals is also responsible for directional differences in solubility observed with many solids. In such solids certain faces of the crystal will dissolve faster than others to produce characteristic patterns on the surface called *etch figures*. These etch figures are quite typical, and when observed carefully under a microscope may be employed to characterize the nature and even composition of the substance. This method of identification and estimation is employed very extensively in metallographic analysis.

Behavior of Crystals in Polarized Light. According to the electromagnetic theory, light is conceived as being a wave motion in which the vibrations occur in a plane perpendicular to the direction of propagation of the wave. In general, these vibrations will be distributed at all angles within this plane. However, when a beam of light is passed through a crystal such as calcite, not only is the beam doubly refracted, but both refracted rays are also polarized. In a polarized beam light does not vibrate any more at all angles in a plane perpendicular to the direction of propagation, but the vibrations are confined to one fixed angle of vibration. The state of affairs in an unpolarized and polarized beam may be visualized from the illustration in Fig. 3.

The two polarized beams of light produced by a calcite crystal are termed the *ordinary* and *extra-ordinary* beams, the angles of vibration of the two being perpendicular to each other. If the calcite crystal is cut in such a manner that the ordinary beam is always completely reflected internally, while the extra-ordinary beam is permitted to pass

out, we have a means of obtaining a single beam of polarized light. Such a beam will not be obstructed by another calcite crystal similarly cut and placed. If, however, the second crystal is rotated about the axis coincident with the direction of light propagation, the light intensity will be found to decrease and will be completely extinguished when an

NORMAL LIGHT RAY
SHOWING VIBRATIONS
AND DIRECTION OF
LIGHT PROPAGATION

POLARIZED LIGHT WAVE
SHOWING SINGLE ANGLE
OF VIBRATION AND
DIRECTION OF LIGHT
PROPAGATION
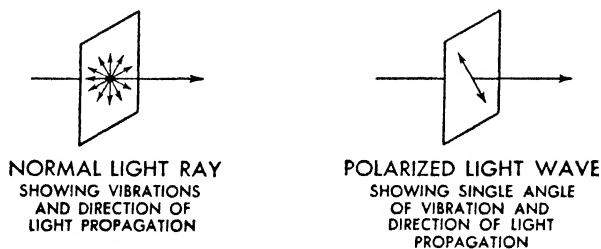
Fig. 3. Polarized and Unpolarized Light

angle of rotation of 90° is reached. The extinction of light occurs because the second crystal, like the first, has been cut to transmit the light at a definite angle and will not transmit light when turned through 90°. At 180° rotation the light will again be completely visible, while at 270° it will again be completely extinguished. Each of these calcite crystals is known as a *Nicol prism*, the first being termed the *polarizer*, the second the *analyzer*.

A combination of Nicol prisms with a microscope is known as a *polariscope* or *polarizing microscope*. Anisotropic crystals examined with a polarizing microscope yield a variety of characteristic and frequently vivid interference patterns which yield valuable information as to the crystalline nature and form of the substances. For this reason the polarizing microscope finds frequent and useful application in mineralogy and chemistry. It cannot be applied in this manner to crystals in the regular system, but only to those which are uni- or biaxial.

Another use of the polarizing microscope is in determining the degree of rotation of the plane of polarized light produced by certain crystals. If for instance a section of quartz of measured thickness is placed between two Nicol prisms and then viewed along the optical axis, it will be found that the maximum intensity does not occur when the Nicols are at the same angle, but at some other definite angle, since the quartz itself has the power of rotating the plane of polarization of the light. In a 1 mm section of quartz this rotation amounts to 22°. For any given wave length of light the degree of rotation is an additive property; i.e., the amount of rotation is proportional to the thickness of section viewed.

Polymorphism. Many substances exist only in a single solid crystalline form. Quite frequently it is found, however, that certain substances

occur in more than one solid modification, or undergo a change of crystalline form on heating or under pressure. The existence of a substance in more than one modification is known as *polymorphism*. Thus carbon exists in crystalline form as either diamond or graphite, calcium carbonate occurs as calcite or aragonite, while sulfur has been found to exist in a variety of solid modifications. Polymorphism occurring in elements, as in the last illustration, is more commonly referred to as *allotropy*. Some polymorphs exist as modifications in the same crystal system, as in ammonium chloride where both forms are cubic, or they may crystallize in different crystal systems, as in rhombic and monoclinic sulfur, or hexagonal and cubic silver iodide.

In general, each polymorphic form of a substance is thermodynamically stable within a particular temperature and pressure range, and the transformation from one form to another takes place at any given pressure at a fixed temperature, known as the *transition temperature* or *transition point*. The rhombic variety of sulfur is stable at atmospheric pressure up to 95.6° C, at which temperature it is transformed into the monoclinic form. Conversely, on cooling slowly monoclinic sulfur, rhombic sulfur will not appear until the temperature of 95.6° C has been reached. At this temperature monoclinic and rhombic sulfur are in equilibrium, and the temperature will not change until *all* the monoclinic sulfur has been transformed to the rhombic variety. In this respect the transition temperatures bear a very striking similarity to fusion points, and are sometimes employed as fixed points in thermometry.

In instances like the preceding the change is entirely reversible, i.e., it proceeds in both directions, and the temperature at which both forms are in equilibrium is definite and constant. However, the change from one form to the other often is not reversible, but proceeds in one direction only. The fact is that if more than one form of a substance is found to exist within the same temperature and pressure range, only one of these is stable. All others are unstable and *tend* to change continuously and irreversibly into the stable form. Nevertheless, the rate of change is often extremely slow, and the unstable form or forms may exhibit to all appearances all the attributes of stability. The existence of certain metastable forms over long geologic periods is ample proof of the sluggishness of some transitions. An instance of such a metastable substance is aragonite ($CaCO_3$), which, although apparently stable, may be transformed into the more stable calcite. However, the reverse transformation, calcite to aragonite, is impossible at ordinary pressures, a fact which speaks for the thermodynamic instability of aragonite.

Table 3 lists several polymorphic substances, the transformations which they undergo, and the transition temperatures for these.

TABLE 3

TRANSITION POINTS OF POLYMORPHIC SUBSTANCES

Substance	Transition	Transition Temp.
Sulfur	Rhombic $\rightleftharpoons$ Monoclinic	95.6°
Tin	Grey $\rightleftharpoons$ White (tetragonal)	18°
	White $\rightleftharpoons$ Rhombic	170°
Ammonium nitrate	Tetragonal $\rightleftharpoons$ (Rhombic)$_1$	$-17°$
	(Rhombic)$_1$ $\rightleftharpoons$ (Rhombic)$_2$	32.1°
	(Rhombic)$_2$ $\rightleftharpoons$ Rhombohedral	84.2°
	Rhombohedral $\rightleftharpoons$ Cubic	125.2°
Potassium nitrate	Rhombic $\rightleftharpoons$ Rhombohedral	128.5°
Silver iodide	Hexagonal $\rightleftharpoons$ Cubic	146.5°
Silver nitrate	Rhombic $\rightleftharpoons$ Rhombohedral	159.5°

Variation of Transition Temperature with Pressure. The transition temperature, like the melting point, is a function of the pressure. For any transformation $A \rightleftharpoons B$ the variation of the transition temperature with pressure is given by the Clapeyron equation,

$$\frac{dT}{dP} = \frac{T(V_B - V_A)}{\Delta H_t} = \frac{T\Delta V}{\Delta H_t} \tag{10}$$

where T is the temperature of transition, P the pressure, ΔV the volume change on transition, and ΔH_t is the heat of transition, defined by

$$\Delta H_t = H_B - H_A$$

Equation (10) is handled in calculations in exactly the same way as the similar equation for the variation of the melting point with pressure, equation (3). ΔH_t is always positive where the transformation is from the form stable at low temperatures to that at high temperature, and negative for the reverse change. It will be observed that for a positive ΔH_t and a positive change in pressure (increase), the sign of ΔT depends on the sign of ΔV. When there is an increase in volume on transition, the transition temperature is raised by an increase in pressure, while when ΔV is negative the transition temperature is lowered.

THE STRUCTURE OF CRYSTALS

X Rays. Although crystallographic study has yielded valuable information concerning the geometrical configuration of crystals, such study does not and cannot tell us what are the structural units of a crystal, and how these units are combined to form the crystal. The answer to

these questions has been made possible only by the discovery of x rays by Roentgen, and by the application of these rays to crystal structure analysis by von Laue, the Braggs, father and son, Debye and Scherrer, Hull, and others. The contribution to the understanding of the solid state which this new mode of attack has made possible stands out in importance and interest as one of the major achievements of twentieth century science.

X rays were discovered by Roentgen in 1895. In that year Roentgen observed that cathode rays (electrons) generated in an evacuated Crookes tube, upon striking the glass walls of the tube, caused it to emit a new and invisible radiation. Since he did not know the nature of these rays, he called them x rays. They were found to be generated more powerfully when the electrons were impinged upon a heavy metal target, such as platinum, tungsten, or molybdenum, instead of on glass. These rays, like visible light, darkened a photographic plate, but the surprising feature about them was that they could do this even when several inches of a metal or some other opaque material were interposed between the plate and the source of the radiation.

X rays, like ordinary light, were subsequently shown to be a form of electromagnetic radiation, but of very much shorter wave length. The
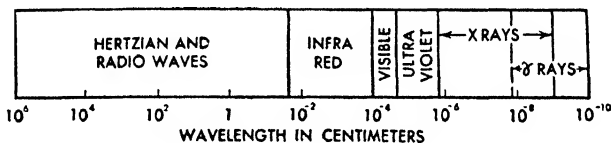
Fig. 4. Spectrum of Electromagnetic Radiation

position of x rays in the electromagnetic radiational spectrum can best be judged from Fig. 4, where the various forms of such radiation are arranged in order of decreasing wave length. It will be observed that the wave lengths of x rays center approximately about 10^{-8} cm, or 1 angstrom (Å), in length, a distance which is of the same order of magnitude as molecular diameters in gases and the roughly estimated interatomic distances in a solid.

The Laue Method of X-Ray Analysis. When x rays were first investigated, the problem arose of measuring their wave length. It is a well-known fact that, if light is permitted to strike a surface consisting either of a series of edges or of lines spaced closely enough to be of the same order of magnitude as that of the wave length of the light, the beam of light is diffracted, and the various radiations are dispersed into a series of spectra known as the first, second, third, etc., order spectra. Furthermore, there is a definite relation between the angle of diffraction, the wave length of the radiation, and the spacing of the lines on the

ruled grating. Since x rays are of the same nature as light, it should be theoretically possible to determine the wave length of this radiation in the same way. However, it is impossible by any mechanical means to rule a grating as fine as that required, namely, one with about 10^8 lines per centimeter. To overcome this difficulty, Max von Laue in 1912 made the brilliant suggestion that, if a crystal consists actually of an orderly arrangement of atoms, then the atomic planes in the crystal should be spaced at intervals of about 10^{-8} cm, and the crystal should
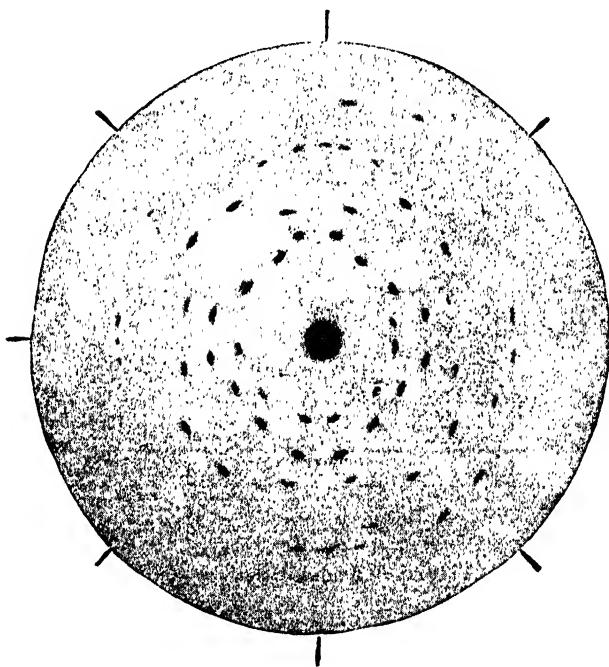
Fig. 5. Laue Diffraction Pattern of Zinc Blende (From Davey, W. P., *A Study of Crystal Structure and Its Applications*, McGraw-Hill Book Company, Inc.)

act then as a natural and very fine three-dimensional diffraction grating for x rays. He showed by a mathematical analysis that, if a beam of x rays were directed against a crystal, the beam should be diffracted and transmitted through the crystal in such a way that the image caught on a photographic plate behind the crystal should show a series of spots arranged in some geometrical fashion about the center of the beam.

Following von Laue's suggestion, Friedrich and Knipping undertook the experimental investigation of the problem in 1913, and showed that Laue's predictions were verified in every respect. They obtained a series

of diffraction patterns for various substances which showed differences characteristic of the substances examined. One of these Laue diagrams, that of zinc blende (ZnS), is shown in Fig. 5. These diagrams speak for a definite arrangement of the atoms in a crystal, and actually permit a reconstruction of the crystal arrangement which would account for the particular distribution of each Laue pattern. However, the method of reconstruction is highly complicated and difficult. A much simpler method of crystal analysis is that suggested by W. H. and W. L. Bragg, and this will be discussed now in some detail.

The Bragg Method of Crystal Analysis. The Braggs first called attention to the fact that, since a crystal may be considered as composed of a series of equally spaced atomic planes, it may be employed not only as a transmission grating, as in the Laue method, but also as a reflection
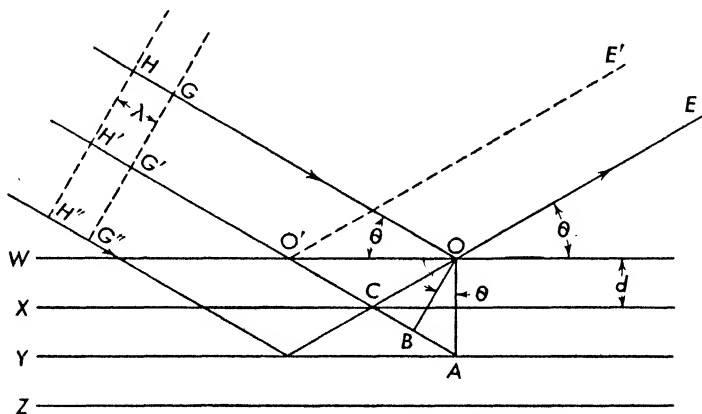
Fig. 6. Reflection of X Rays by Parallel Planes

grating. A beam of x rays striking the atoms which constitute these planes will be diffracted then in such a manner as either to cause interference with or reinforcement of the beam diffracted from the first, or outer, plane, and the whole beam will behave as if it had been *reflected* from the surface of the crystal.

To understand better the theory of this method consider, as shown in Fig. 6, a wave front $GG'G''$ of x rays approaching at an angle θ a series of parallel, equidistant planes W, X, Y, Z, etc., which constitute the atomic planes of the crystal. Part of the beam HGO will be reflected at O along OE at the angle of reflection θ, which is the same as the angle of incidence. Similarly the beam $H'G'O'$ will be reflected partly at O' along $O'E'$, and then again at C on the second plane along COE. In order to emerge along OE the second beam has to travel a longer distance than the first, namely, the distance $H'G'O'CO$ as against the distance HGO. It is an established fact that if the difference in distance of the

two paths is exactly equal to an integral multiple of the wave length of the radiation, the two beams will be in phase at O, will reinforce each other, and the intensity of the reflected rays will be at a maximum. When the two beams are out of phase, however, interference will result, and the intensity of the reflected beam will be less than the maximum. The condition for maximum reflection intensity is then that the distance

$$G'CO - GO = n\lambda$$

where λ is the wave length of the x rays employed, while n is an integer taking on the values $1, 2, 3, \cdots$ and known as the *order of reflection*.

If a perpendicular is drawn now from O to the extension of the line $G'O'C$, while another is dropped from O to A perpendicular to W, X, and Y, it follows that

$$G'CO = G'A - CA + CO$$
$$= G'A$$

since $CO = CA$ from the construction of the figure. But

$$GO = G'B$$

Therefore, $\quad G'CO - GO = G'A - G'B$
$$= BA$$

and hence, $\quad\quad\quad BA = n\lambda$

It can readily be shown that the angle BOA is also θ. Then, since OB was constructed perpendicular to $G'A$,

$$\sin \theta = \frac{BA}{OA}$$

and $\quad\quad\quad\quad BA = OA \sin \theta$
$$= 2\,d \sin \theta$$

where d is the distance between any two atomic planes in the crystal.

Therefore, $\quad\quad\quad n\lambda = 2\,d \sin \theta \quad\quad\quad\quad\quad (11)$

This simple equation connects directly the wave length and order of reflection of the x rays with the interplanar distance d and the angle of maximum reflection θ. Without any further information the ratio λ/d is deducible by measuring n and θ. On the other hand, if λ is known somehow, d may be calculated; or, vice versa, if d is known, the crystal may be employed to determine the wave length of the x rays.

The reflection angles θ and the intensities of the reflected beams corresponding to these angles are determined with the Bragg x-ray spectrometer, a diagram of which is shown in Fig. 7. The x rays generated in tube A by bombardment of a suitable target B are passed through a series of slits and screen (C, D, E) to give a sharp and monochromatic

beam, and are then directed to strike the face of a crystal which is suitably mounted on a turntable F. The graduated turntable may be rotated to give any angle of incidence desired. Coaxially with the table and crystal is mounted an ionization chamber H, into which the reflected beam is passed. The ionization of the gas filling the chamber, usually sulfur dioxide, is proportional to the intensity of the x rays passing
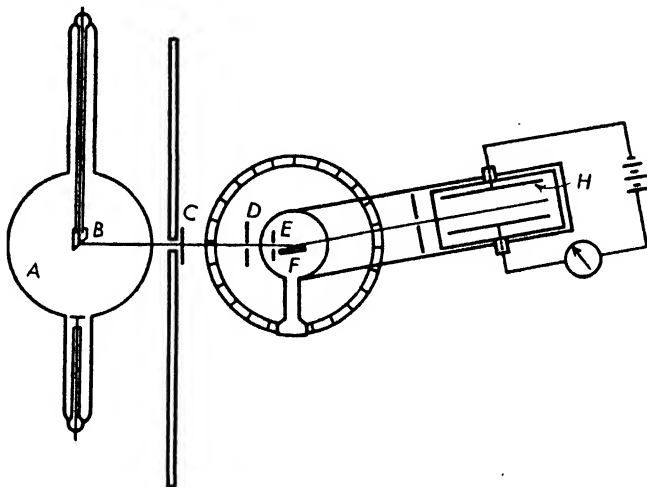
Fig. 7. Bragg X-ray Spectrometer

through the chamber. Since the current passing through the ionization chamber is proportional to the ionization of the gas, the intensity of such a current as measured by an electrometer gives directly a measure of the relative intensities of the x rays that are reflected from the crystal. By determining the intensities of the reflected beam at various angles of reflection, the angles at which maximum reflection occurs may be readily found.

The X-Ray Analysis of Sodium Chloride. The procedure employed by the Braggs to elucidate the structure of sodium chloride will now be described in some detail, and this example may be taken as a general indication of the methods involved in studying any other type of crystal. Sodium chloride has been chosen because it belongs to the cubic system, exhibits the highest type of symmetry in crystals, and is the simplest to study.

Any macroscopic crystal of a substance may be thought of as built up by the repetition of a fundamental structural unit known as a *unit crystal lattice*. Each unit lattice, in turn, must be constituted of atoms, molecules, or ions, as the case may be, arranged to give the particular geometrical configuration of the lattice. The unit lattice of sodium chloride,

like the macroscopic crystal, must be a cube, and atoms of sodium and chlorine must be arranged so as to give this cube. To meet the requirements of cubic symmetry the atoms within the cube can only be arranged in one of three possible ways, known respectively as (a) the *simple cubic*, (b) the *face-centered cubic*, and (c) the *body-centered cubic* arrangements. These are illustrated in Fig. 8. In the simple cubic lattice an atom is located at each of the corners of the cube. The face-centered cubic lattice in turn involves a simple cubic arrangement modified by the location of an atom in the center of each of the six faces of the cube. Finally, the body-centered cubic lattice is again a simple cubic arrangement but modified this time by the presence of a single atom in the *center* of the cube.

In the cubic system all the possible planes that can be passed through the atoms have the Miller indices (100), (110), or (111), as may be seen from Fig. 9. However, the ratios of the distances among the (100), (110), and (111) planes in the three types of cubic lattices are not the same. If we arbitrarily designate by a the distance between 100 planes in the simple cubic lattice, then the perpendicular distance between (110) planes will be $a/\sqrt{2}$, and the distance between (111) planes will be $a/\sqrt{3}$. In the face-centered cubic lattice, on the other hand, parallel

SIMPLE CUBIC LATTICE

BODY-CENTERED CUBIC LATTICE

FACE-CENTERED CUBIC LATTICE
Fig. 8

planes can be interposed halfway between the (100) and the (110) planes in the simple cubic lattice, and hence the interplanar distances in this case are respectively $a/2$, $a/(2\sqrt{2})$, and $a/\sqrt{3}$. Finally, in the body-centered cubic lattice parallel planes can be interposed halfway between any (100) or (111) planes in the simple cubic lattices, so that in terms of a the respective distances become $a/2$, $a/\sqrt{2}$, and $a/(2\sqrt{3})$. These distances for the various planes in the three types of cubic lattices are summarized in Table 4. The ratios $d_{100} : d_{110} : d_{111}$ for the simple cubic, face-centered cubic, and body-centered cubic lattices are then:

$$\text{Simple cubic: } d_{100} : d_{110} : d_{111} = a : \frac{a}{\sqrt{2}} : \frac{a}{\sqrt{3}} = 1 : 0.707 : 0.577$$

$$\text{Face-centered cubic: } d_{100} : d_{110} : d_{111} = \frac{a}{2} : \frac{a}{2\sqrt{2}} : \frac{a}{\sqrt{3}} = 1 : 0.707 : 1.154$$

$$\text{Body-centered cubic: } d_{100} : d_{110} : d_{111} = \frac{a}{2} : \frac{a}{\sqrt{2}} : \frac{a}{2\sqrt{3}} = 1 : 1.414 : 0.577$$

TABLE 4

DISTANCE BETWEEN ATOMIC PLANES FOR CUBIC CRYSTALS

	(100)	(110)	(111)
Simple cubic	a	$\dfrac{a}{\sqrt{2}}$	$\dfrac{a}{\sqrt{3}}$
Face-centered cubic	$\dfrac{a}{2}$	$\dfrac{a}{2\sqrt{2}}$	$\dfrac{a}{\sqrt{3}}$
Body-centered cubic	$\dfrac{a}{2}$	$\dfrac{a}{\sqrt{2}}$	$\dfrac{a}{2\sqrt{3}}$

It will be observed that the ratios of the distances in the three cases are all different, and hence a determination of the interplanar distance ratios in sodium chloride should permit a decision as to the type of lattice
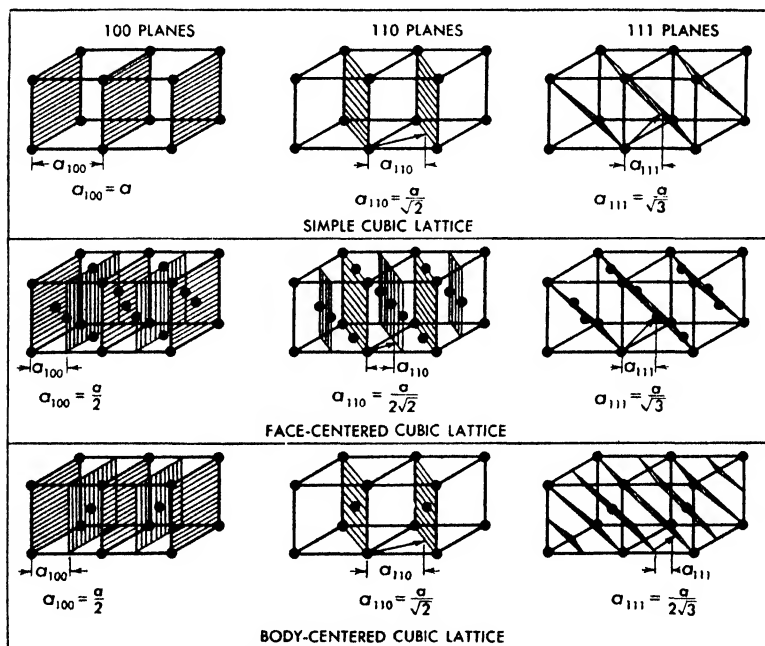
Fig. 9. Planes in Cubic Lattices

to which this substance belongs. Since for any given reflection order and wave length of x rays $d = (n\lambda)/(2 \sin \theta)$, then for nth order reflection maxima from the (100), (110), and (111) planes we have

$$d_{100} : d_{110} : d_{111} = \frac{n\lambda}{2 \sin \theta_1} \cdot \frac{n\lambda}{2 \sin \theta_2} \cdot \frac{n\lambda}{2 \sin \theta_3}$$

$$= \frac{1}{\sin \theta_1} : \frac{1}{\sin \theta_2} : \frac{1}{\sin \theta_3}$$

and a knowledge of the angles at which maximum intensities of reflection occur for the three types of planes is sufficient to determine the type of lattice to which sodium chloride belongs. In Table 5 are given the values of θ, the angle of maximum reflection, for several values of n for each of the three types of planes, as well as sin θ. Considering first order reflections only, we see that

$$d_{100} : d_{110} : d_{111} = \frac{1}{\sin \theta_1} : \frac{1}{\sin \theta_2} : \frac{1}{\sin \theta_3} = \frac{1}{0.103} : \frac{1}{0.146} : \frac{1}{0.0906}$$
$$= 1 : 0.705 : 1.14$$

Comparing these ratios with those previously established for the three possible types of arrangement of cubic lattices, it follows that in sodium chloride the atoms must be arranged in a face-centered cubic lattice. The same conclusion may be arrived at by considering the results for second order reflection maxima.

TABLE 5

GLANCING ANGLES FOR NaCl USING K LINE FROM PALLADIUM

Planes	First Order		Second Order		Third Order	
	θ_1	Sin θ_1	θ_2	Sin θ_2	θ_3	Sin θ_3
(100)	5.9	0.103	11.9	0.208	18.2	0.312
(110)	8.4	0.146	17.0	0.292	—	—
(111)	5.2	0.0906	10.5	0.182	—	—

The Relative Positions of Sodium and Chlorine Atoms in Lattice. Determining the type of lattice along which the atoms in sodium chloride are arranged does not solve completely the problem of the structure of sodium chloride, for the question of the relative arrangement of the atoms of sodium and chlorine in the lattice still remains. This question can be resolved only by considering the relative intensities of the reflection maxima for the different orders and planes. In Table 6 are given the intensities for the various orders of reflection for the (100), (110), and (111) planes, the intensity of the first order maximum for the (100) plane being taken as 100.

It is an established fact that the intensity of a diffracted beam depends, in the first place, on the mass of the particle responsible for the diffraction, being greater the larger the mass; and, second, on the order of diffraction, the intensity decreasing in a definite manner with increase in order. These two facts are employed to explain the results of Table 6.

TABLE 6

RELATIVE INTENSITIES FOR REFLECTION MAXIMA IN NaCl

Order	Intensities for (100) Planes	Intensities for (110) Planes	Intensities for (111) Planes
First	100.00	50.4	9.00
Second	19.90	6.10	33.1
Third	4.87	0.71	0.58
Fourth	0.79		2.82
Fifth	0.12		0.14

It will be observed that in the (100) and (110) planes the intensities decrease progressively with order. This systematic decrease can be accounted for quantitatively by assuming that such planes contain equal numbers of sodium and chlorine atoms. In the (111) planes, however, an alternation of intensities is observed, the first order being weaker than the second and the third weaker than both first and second, while the fourth is weaker than the second, but stronger than the third. It is possible to account for this by postulating that the (111) planes are composed alternately of sodium atoms and chlorine atoms, and that the planes containing sodium atoms only are interposed halfway between the (111) planes containing only chlorine atoms.
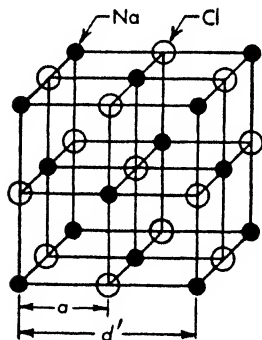
Fig. 10. The Sodium Chloride Lattice

Exhaustive study has shown that the only possible arrangement of sodium and chlorine atoms which will satisfy the facts elicited from the study of the intensities is that shown in Fig. 10. In this arrangement atoms of sodium, shown by black circles, are located in the corners of the cube and in the center of each of the six faces. Halfway between every two sodium atoms is located a chlorine atom, shown by open circles. A moment's reflection will show that the indicated structure meets the required conditions, and that it consists essentially of two interpenetrating face-centered cubic lattices, one composed entirely of sodium atoms, the other of chlorine atoms. The chlorine lattice is merely displaced the distance a along any edge of the cube.

Although for purposes of discussion we have considered the structural units of sodium chloride to be sodium and chlorine atoms, the general

consensus at present seems to be that the sodium chloride in the crystal is ionized, and that the units are rather sodium and chloride *ions*. Again, in terms of the x-ray interpretation of the structure as outlined here, the designation "molecule of sodium chloride" loses a great deal of its definite meaning. It can hardly be said that any particular chloride ion belongs to any definite sodium ion; rather, each sodium is shared equally by six chloride ions, and each chloride is held equally by six sodium ions. All that can be said is that to each sodium corresponds *one-sixth* of *six* chloride ions, so that each sodium has the equivalent of a chloride, but not any one ion exclusively.

Calculation of *d* and λ. The cube indicated in Fig. 10 is considered to be the unit crystal, or unit lattice, of sodium chloride, and the edge of the cube is then the distance between any two sodium or any two chloride ions. This is *twice* the distance between (100) planes. The macroscopic crystal is built up by the extension of unit lattices in all directions.

Once the nature of the unit lattice is established, it is a simple matter to calculate the length of the edge of the unit cube and the distance between (100) planes from the molar volume and Avogadro's number. Each of the sodium ions at the corners is shared by eight cubes, hence to each cube may be ascribed 1/8 (8), or one sodium ion. Further, each of the sodium ions in the centers of the faces is shared by two cubes. This adds 1/2 (6) or three sodium ions more, making a total of four sodium ions. Similarly each of the chloride ions along the edges is shared by four cubes, and since there are 12 of these, 1/4 (12) or three chloride ions are part of the cube. Adding to these the one in the center of the cube, we obtain $3 + 1 = 4$ chloride ions as the average chloride content of the cube. It may be said, therefore, that each unit lattice of sodium chloride contains on the average four sodium ions and four chloride ions, or a *total equivalent to four molecules of sodium chloride.*

The molar volume of sodium chloride is the gram molecular weight, 58.45, divided by the density, 2.17 gm/cubic centimeter, or

$$V_m = \frac{58.45}{2.17} = 26.93 \text{ cc}$$

This is the volume occupied by $N = 6.02 \times 10^{23}$ molecules. The volume occupied by four molecules, V, which is also the volume of the unit cube, is, then,

$$V = \frac{V_m \times 4}{6.02 \times 10^{23}} = \frac{26.93 \times 4}{6.02 \times 10^{23}} = 178.9 \times 10^{-24} \text{ cc}$$

From this volume the edge of the unit cube follows as

$$d' = \sqrt[3]{178.9 \times 10^{-24}}$$
$$= 5.63 \times 10^{-8} \text{ cm}$$

Since the edge of the unit cube is twice the distance between (100) planes, the distance $d = a$ for sodium chloride to be employed in the Bragg equation is $d'/2$, or,

$$d = \frac{5.63 \times 10^{-8}}{2} = 2.815 \times 10^{-8} \text{ cm}$$

The best value accepted at present for this distance is $2.8140 \pm 0.0010 \times 10^{-8}$ cm.

Once the interplanar distance for sodium chloride is known, the crystal may be employed to determine the wave length of any x rays which may be directed against it. Thus in the Bragg experiments $\sin \theta$ was found to be 0.103 for $n = 1$. Consequently the x rays employed in these experiments had a wave length of

$$\begin{aligned}
\lambda &= \frac{2\,d \sin \theta}{n} \\
&= \frac{2(2.814 \times 10^{-8})\, 0.103}{1} \\
&= 0.58 \times 10^{-8} \text{ cm} \\
&= 0.58 \text{ Å}
\end{aligned}$$

The Debye-Scherrer-Hull Method of Crystal Analysis. In many instances it is impossible to obtain a macroscopic crystal of a substance suitable for the Bragg method of analysis. To obviate this difficulty, Debye and Scherrer in 1916, and independently Hull in 1917, devised a method of x-ray crystal analysis which permits the use of a substance in *powder* form. In this method a monochromatic beam of x rays is focused upon a small tube containing the finely ground substance to be examined. Since in the powder the crystal planes are oriented at all possible angles to the beam, there will be always some crystals with just the proper orientation to give reinforced diffraction images, and from all the planes simultaneously. These diffraction maxima are photographed on a film fixed behind the sample in the form of a circular arc.

The diffracted rays obtained in this manner form concentric cones originating from the powder under examination. Photographed on a narrow strip of film, these images appear as nearly vertical lines arranged on each side of a bright center spot due to the undiffracted beam, although actually these lines are portions of arcs of circles whose center is the point of focus of the x-ray beam. Each pair of lines equidistant to the right and left of the center spot corresponds to a single order of diffraction for a family of planes, a pair being obtained for each order of diffraction for each type of plane present.

A powder diffraction pattern of magnesium oxide is shown in Fig. 11.

The diffraction patterns obtained from various solid substances all are characteristic of the substances responsible for the particular patterns. For this reason the Debye-Scherrer-Hull method of x-ray diffraction is frequently employed in qualitative and quantitative chemical analysis to identify and estimate both pure substances and mixtures. When used for analytical purposes a comparison plate is prepared showing the position of various lines for the various substances which may be expected
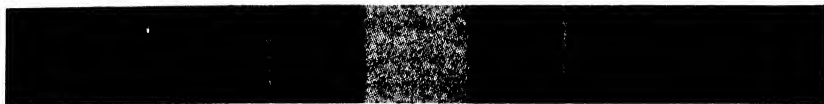
Fig. 11. Powder Diffraction Diagram of MgO (From Wyckoff, R. W. G., *The Structure of Crystals*, Reinhold Publishing Corporation)

in the unknown sample. Then several milligrams of the unknown sample are also exposed, and the diffraction plate thus obtained is compared with the reference plate. Occurrence of identical lines in both plates testifies to the presence of the particular substance exhibiting those lines in both samples. The identity of each line is determined by comparison of plates with diagrams obtained from the pure constituents.

For quantitative estimation of the amount of each substance present the intensity of each line must be determined and compared with the intensities obtained from definite amounts of the pure constituents under identical operating conditions. Since the intensity of a line is proportional to the amount of substance present, such a comparison gives directly an estimate of amount. In the more elaborate installations intensities are evaluated with photoelectric comparators.

In one respect an x-ray analysis supplies more information than a chemical analysis. From a qualitative chemical analysis of a mixture of, say, calcium chloride and sodium bromide, it is impossible to tell whether the constituents are calcium chloride and sodium bromide or calcium bromide and sodium chloride. The x-ray method answers this question directly, for it will show the lines of calcium chloride and sodium bromide, and no lines for the reciprocal salts.

X rays are also employed extensively in medicine for diagnosis of disease. In metallurgy and industry x rays are utilized to detect structural flaws in metals, as well as to identify and control the impurities, both desirable and undesirable, which occur. If anything, with the development of more powerful and penetrating equipment, the future promises even a wider application of these rays to various theoretical and technical problems.

Results of X-Ray Study of Crystals. Studies similar to that described for sodium chloride have been made upon a great many solid

substances to determine their structure. Some of the results of these studies will now be summarized.

Many metals crystallize in the cubic system, with the atoms arranged in face-centered or body-centered lattices. As examples of the first may be cited aluminum, calcium, nickel, cobalt, copper, silver, platinum, gold, and lead, while of the second lithium, chromium, sodium, potassium, iron, and tungsten. The simple cubic arrangement does not appear to be overly favored. On the other hand, the sodium chloride structure of two interpenetrating face-centered lattices is very common among compounds falling in the regular system. It is exhibited by oxides such as those of magnesium, calcium, strontium, barium, nickel, and cobalt, sulfides such as those of magnesium, barium, manganese, and lead, and all of the halides of the alkali metals except those of cesium. The cesium halides occur as two simple cubic lattices, one of cesium and another of the halide, interlocking to form a resultant body-centered lattice with the equivalent of one molecule of cesium halide per unit cube. This structure is shown in Fig. 12.

● Cs
○ Cl

Fig. 12. The Cesium Chloride Lattice

Another type of cubic arrangement is found in zinc sulfide, diamond, silicon, germanium, and gray tin. Here we have a simple face-centered lattice which also contains an atom in the center of each alternate small cube within the larger lattice. On this basis each atom is equidistant from four other atoms. A modification of this scheme is the arrangement found in calcium, strontium, and barium fluorides. In these salts the metal ions are located on a face-centered lattice, while the nonmetal ions are situated in the centers of each of the eight small cubes composing the metallic lattice. The nonmetallic ions thus form a simple cubic lattice inside the face-centered lattice of the metal ions.

A still different cubic structure is that shown by cuprous oxide and silver oxide, Fig. 13. In crystals of these the oxygens lie on a body-centered cubic lattice, while the metal atoms interpenetrate this lattice to fall at the center of each alternate small cube within the oxygen framework.

○ O
○ Cu
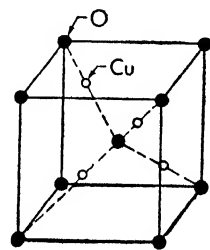
Fig. 13. The Cuprous Oxide Lattice

As an example of a noncubic arrangement may be given the close-packed hexagonal lattice, Fig. 14, which occurs in elements such as magnesium, zinc, cadmium, and titanium. The compounds zinc oxide, beryllium oxide, cobalt sulfide, and stannous sulfide appear as two such

interpenetrating hexagonal lattices, one composed of the metal, the other of the nonmetal. On the other hand, calcium, magnesium, manganese, and iron carbonates and sodium nitrate have a rhombohedral lattice in which the axes are all equal but inclined at equal angles other than 90°. If we imagine this lattice to be a distorted cube, then the metal ions and the anions are on two interpenetrating distorted face-centered lattices comparable to the undistorted sodium chloride lattice. It is of interest to point out that the anions CO_3^{--} and NO_3^- have been shown to be present in the crystal as a unit, with the oxygens distributed about the central atom. The same is true of other inorganic anions like SO_4^{--}, PO_4^{---}, ClO_3^-, and MnO_4^-. These findings are excellent evidence for the dissociation of salts in the solid state and for the existence of these radicals as individual entities.
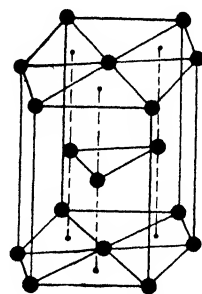
Fig. 14. Close-Packed Hexagonal Arrangement

Lattice Structural Units. The units entering into the construction of a crystal lattice may be ions, atoms, or molecules, depending on the nature of the substance in question. Inorganic compounds are considered generally to have ionic lattices, like sodium chloride. In such lattices the binding forces are the electrostatic attractions between the oppositely charged ions. As has already been pointed out, for such substances the ordinary definition of a molecule is meaningless, for no particular negative ion, say, may be said to belong exclusively to any particular positive ion. Rather, because of the requirements of electroneutrality, each positive ion has associated with it on the average a sufficient number of negative charges to balance the positive charge. Ionic lattices are generally very stable, and crystals constituted of these melt at relatively high temperatures.

In metallic crystals and in diamond and graphite the lattice-building units are atoms. Within the diamond lattice each atom of carbon is surrounded by four others in the form of a regular tetrahedron; this arrangement confirms the valency of carbon and the directionality of the valence bonds as postulated by the organic chemists. In graphite the structure consists of connected planar hexagonal rings linked vertically to similar planar rings to produce the characteristic layer arrangement of graphite. For the metals some of the structures have already been mentioned.

In substances like carbon dioxide, hydrogen chloride, and stannic iodide the complete molecule acts as a structural unit within the lattice. X-ray evidence indicates that such molecules occupy the key spatial positions, while the atoms within the molecules are arranged in a definite

configuration about the mean position of equilibrium for each molecule. The molecules are held in their geometrical distribution within the lattice by van der Waals forces. These forces are much weaker than the electrostatic attractions in ionic lattices, and consequently such crystals are less strong and melt at considerably lower temperatures. The lower rigidity and higher vapor pressure of these substances as compared with ionic crystals may also be ascribed to the relative weakness of the van der Waals forces as compared to electrostatic attractions.

Electron Diffraction Method of Crystal Analysis. DeBroglie in 1924 first called attention to the fact that moving electrons, besides exhibiting their corpuscular properties, should also possess properties characteristic of wave motion of the type associated with light. He showed that an electron moving with velocity v should have associated with it a wave motion of wave length λ given by

$$\lambda = \frac{h}{mv} \tag{12}$$

where m is the mass of the electron, 9.1×10^{-28} grams, while h is Planck's constant, namely, 6.62×10^{-27} erg-sec. The velocity of an electron depends on the potential drop through which the electron falls. For the potential drop, $\mathcal{E}$, in volts it is given as

$$v = 5.94 \times 10^7 \sqrt{\mathcal{E}} \tag{13}$$

Inserting equation (13) in equation (12), we obtain for the wave length, λ,

$$\begin{aligned}
\lambda &= \frac{h}{m \times 5.94 \times 10^7 \sqrt{\mathcal{E}}} \\
&= \frac{6.62 \times 10^{-27}}{9.1 \times 10^{-28} \times 5.94 \times 10^7 \sqrt{\mathcal{E}}} \\
&= \frac{12.2 \times 10^{-8}}{\sqrt{\mathcal{E}}}
\end{aligned} \tag{14}$$

For potentials between 10 and 10,000 volts, λ should vary, then, between 3.86 and 0.12 Å, and hence such electrons should behave like x rays toward crystals.

This prediction of DeBroglie's was confirmed in a beautiful manner by Davisson and Germer in 1925. Davisson and Germer investigated the diffraction of electrons by a nickel surface and found that the electron diffraction pattern thus obtained was very similar to the one obtained with x rays. Further, the wave length of the electrons calculated from the diffraction pattern agreed remarkably well with that calculated from the DeBroglie equation.

These experiments and others carried out since 1925 not only provide excellent confirmation of DeBroglie's theory, but also make available a new tool for investigation of solids. Compared to x rays, electron beams are much less penetrating. Consequently, whereas x rays are diffracted by atomic planes deep within a crystal, electrons are diffracted primarily by planes lying near the surface. Electron diffraction offers, then, a convenient means of investigating the nature of solid surfaces and surface films, and a great deal of work in this field is directed thus toward an elucidation of the nature of oxide and other surface films on solids.

REFERENCES FOR FURTHER READING

1. W. Bragg, "Introduction to Crystal Analysis," George Bell & Sons, Ltd., London, 1928.
2. Bragg and Bragg, "The Crystalline State," The Macmillan Company, New York, 1934.
3. W. P. Davey, "A Study of Crystal Structure and Its Applications," McGraw-Hill Book Company, Inc., New York, 1934.
4. J. N. Friend, "A Textbook of Physical Chemistry," J. B. Lippincott Company, Philadelphia, 1933.
5. Kraus, Hunt, and Ramsdall, "Mineralogy," third edition, McGraw-Hill Book Company, Inc., New York, 1936.
6. C. W. Stillwell, "Crystal Chemistry," McGraw-Hill Book Company, Inc., New York, 1938.
7. A. E. H. Tutton, "Crystalline Form and Chemical Constitution," The Macmillan Company, New York, 1926.
8. R. W. G. Wyckoff, "The Structure of Crystals," second edition, Chemical Catalog Company, Inc., New York, 1931.

PROBLEMS

1. The densities of liquid and solid Hg are respectively 13.70 and 14.19 g/cubic centimeter at the melting point, $- 38.87°$ C. For this substance the heat of fusion is 557 cal/gram atom. Find the change in melting point per atmosphere change in pressure.   *Ans.* 0.0051° C/atm

2. Calculate and compare the rates of change of vapor pressure with temperature for water and ice at 0° C.

3. The vapor pressure of $C_6H_6(s)$ is 5.79 mm at $- 20°$ C, while that of $C_6H_6(l)$ is 118.24 mm at 30° C. The heats of fusion and vaporization are, respectively, 30.1 and 107 cal/gram. From these data find (a) the heat of sublimation and (b) the melting point of C_6H_6.   *Ans.* (a) 10,700 cal/mole; (b) 4.3° C

4. The vapor pressure of solid CO_2 is 76.7 mm at $- 103°$ C, and 1 atm at $- 78.5°$ C. Calculate the heat of sublimation of CO_2.

5. NH_4NO_3 undergoes a transition from one solid modification to another at 125.5° C at 1 atm, and at 135.0° C at 1000 atm pressure. The form stable at higher temperatures has an average volume of 0.0126 cc/gram greater than the other modification over the pressure range studied. From these data estimate the heat of transition.

6. The vapor pressure of As(l) may be calculated from the equation

$$\log_{10} P_{mm} = \frac{-2460}{T} + 6.69$$

that of As(s) from the equation

$$\log_{10} P_{mm} = \frac{-6947}{T} + 10.8$$

Find the temperature at which the two forms of As have the same vapor pressure. What is the value of this pressure?

7. The angle of first order reflection of a beam of x rays from a given crystal is 5° 15′. What will be the angle of second order reflection? *Ans.* 10° 33′

8. The first order reflection of a beam of x rays from the (100) face of NaCl occurs at an angle of 6° 30′. What is the wave length of the x rays used?

9. The first order reflections from the (100), (110), and (111) planes of a given cubic crystal occur at angles of 7° 10′, 10° 12′, and 12° 30′ respectively. To what type of cubical lattice does the crystal belong?

10. Al crystallizes with a face-centered cubic lattice. The density of Al is 2.70 g/cubic centimeter at 20° C. Calculate the distance between successive (100) planes in Al. What is the distance of closest approach of Al atoms in the crystal? *Ans.* 2.86 Å

11. The density of CaF_2 is 3.180 g/cubic centimeter at 20° C. With the help of Fig. 14, calculate the dimensions of a unit cube containing 4 Ca^{++} and 8 F^- ions. *Ans.* 5.45 Å

12. Electrons emitted from a hot filament are accelerated in an electric field until their velocity is 5×10^9 cm/second. Using the DeBroglie relation, find the wave length of the electrons.

Solutions

Introduction. When several substances are mixed, three possible types of mixtures may be obtained: (1) a coarse mixture, such as that of salt and sugar; (2) a colloidal dispersion, such as results when fine clay is shaken with water; (3) a true solution, obtained when a substance like sugar dissolves in water. In the coarse mixture, the individual particles are readily discernible, and may be separated from each other by mechanical means. Although in a colloidal dispersion the particles are much finer and the heterogeneity is not so readily apparent, the dispersion is, nevertheless, not homogeneous, as will be shown in the chapter on colloids. On the other hand, in the true solution no heterogeneity can be detected, and the constituents cannot be separated from each other by mechanical means. Every part of the solution is found to be like every other part, i.e., a true solution constitutes a homogeneous phase.

We may, therefore, define a true solution as *a physically homogeneous mixture of two or more substances.* This definition of a solution places no restriction on either the state of aggregation or the relative amounts of the constituents, and consequently a solution may be gaseous, liquid, or solid, and may vary in composition within wide limits. It is this latter fact which excludes pure compounds from the classification of solutions, for a fixed and definite ratio persists among the constituents in a compound.

It is frequently convenient to refer to the substance that dissolves as the *solute,* and to the substance in which solution takes place as the *solvent.* For the solubility of solids in liquids, where the liquid is usually present in large excess over the solid, there is no ambiguity in these terms, the solid being the solute, the liquid the solvent. However, when dealing with the solubility of such liquids as acetone and water or dioxane and water, which dissolve in each other in all proportions, it is difficult to differentiate between solute and solvent. This difficulty will not arise in our discussions, for the terms will be employed only when there is no ambiguity as to meaning.

A solution which contains at a given temperature as much solute as it can hold in presence of the dissolving substance is said to be *saturated.*

Any solution which contains less than this amount of solute is said to be *unsaturated*, while if it contains more than this amount, it is *supersaturated*. A supersaturated solution can exist only in the absence of dissolving substance, and is at best very unstable. Jarring and stirring may, and introduction of solute will, cause the precipitation of excess solute in solution, leading to the formation of a saturated solution. To determine the state of a solution with respect to saturation it is only necessary to introduce some of the dissolving substance. If the substance dissolves, the solution is unsaturated; if no further solubility takes place, the solution is saturated; while if precipitation takes place, the original solution was supersaturated.

Factors Affecting Solubility. The extent to which a substance will dissolve in another varies greatly with different substances, and depends on the nature of the solute and solvent, the temperature, and the pressure. In general the effect of pressure on solubility is small unless gases are involved. However, the effect of temperature is usually very pronounced. The direction in which the solubility of a substance in a solvent changes with temperature depends on the heat of solution. If a substance dissolves at saturation with evolution of heat, the solubility decreases with temperature. On the other hand, if a substance dissolves with absorption of heat, the solubility increases as the temperature is raised.

Our knowledge of intermolecular forces is at present too meager to foretell quantitatively the extent of solubility of one substance in another as a function of the chemical nature of solute and solvent. All that can be said is that in general compounds of similar chemical character are more readily soluble in each other than are those whose chemical character is entirely different. When a similarity of chemical nature exists between two substances, the solution of the two will have an environment which will not be too different from that of the pure substances, and the two can tolerate each other in solution. On the other hand, when the chemical nature of the two substances is considerably different, the two substances cannot tolerate each other, and hence there is little tendency to dissolve. Between these two extremes there may be a considerable number of intermediate stages of similarity which will account for the wide ranges of solubility of various substances in each other.

These points may be illustrated with the phenomena encountered in the mutual solubility of liquids. When ethyl alcohol and water, which are closely related chemically, are mixed, the two dissolve in each other in all proportions, i.e., there is no saturation limit. Such substances are said to be *completely miscible*. In distinction to these, two liquids such as water and mercury, which are very different chemically, do not dissolve in each other at all, and are said to be *completely immiscible*.

Between these two limiting types there are liquid pairs, such as ether and water, which dissolve in each other to a limited extent only. Thus, pure ether dissolves a certain amount of water to form a saturated solution of water in ether, while water dissolves a limited amount of ether to form a saturated solution of ether in water. Consequently, with high proportions of one or the other of these liquids, a completely miscible solution can be obtained. When the proportions taken are outside these saturation limits, however, two layers are obtained, one composed of a solution of ether in water, the other of water in ether. Liquid pairs of this sort are said to be *partially miscible*.

Concentration of Solutions. The concentration of the constituents of a solution can be expressed in many different ways. The following list enumerates the more common of these:

(1) Per cent by weight.
(2) Per cent by volume.
(3) Weight of solute per definite weight of solvent.
(4) Weight of solute per definite weight of solution.
(5) Molarity — number of moles of solute per liter of solution.
(6) Normality — number of equivalents of solute per liter of solution.
(7) Molality — number of moles of solute per *1000 grams of solvent*.
(8) Mol fraction.

All these schemes are either self-explanatory or familiar to the student. As in gas mixtures,[1] the mol fraction of any constituent of a solution is defined as the number of moles of the particular substance present divided by the *total* number of moles of all constituents of the solution. If instead of the number of moles the weights of the constituents are given, the mol fractions can be calculated provided the molecular weights are known, for

$$N_A = \frac{n_A}{n_A + n_B + \cdots} = \frac{\dfrac{W_A}{M_A}}{\dfrac{W_A}{M_A} + \dfrac{W_B}{M_B} + \cdots} \qquad (1)$$

where the W's are the weights of the various species and the M's the respective molecular weights.

The choice of a particular method for expressing concentrations depends entirely on convenience and the purpose at hand. Of the various methods listed above, those expressed on a weight basis, namely, the first, third, fourth, seventh, and eighth, are temperature independent, i.e., the concentrations will be the same at all temperatures. The concentrations expressed on a volume basis, however, such as percentage

by volume, normality, and molarity, will vary with temperature in a manner dependent upon the thermal expansion of the volume of solution.

Types of Solutions. For theoretical purposes it is convenient to subdivide solutions into two major classifications, (1) solutions of nonelectrolytes, and (2) solutions of electrolytes. In nonelectrolytic solutions the solute dissolved in the solvent persists in molecular, uncharged form, and exhibits no tendency to dissociate into electrically charged ions. For such solutions certain general laws have been developed which will be discussed in the following chapter. In electrolytic solutions, on the other hand, the solute dissociates to a greater or lesser degree into ions, increasing thereby the total number of particles in solution. The behavior of the solution with respect to certain properties, therefore, is changed, and the simple laws deduced for nonelectrolytic solutions require modification. For this reason, solutions of electrolytes will be discussed separately in Chapter VII. Certain aspects of solutions, however, can be discussed independently of the above classification. These will be developed in the present chapter.

Although solutions with many components can be prepared, attention will be confined to binary solutions, i.e., solutions containing two components only. Since the solvent and solute may be either gaseous, liquid, or solid, the number of possible types of binary solutions that may be expected is nine, namely:

(1) Solution of a gas in a gas.
(2) Solution of a liquid in a gas.
(3) Solution of a solid in a gas.
(4) Solution of a gas in a solid.
(5) Solution of a liquid in a solid.
(6) Solution of a solid in a solid.
(7) Solution of a gas in a liquid.
(8) Solution of a solid in a liquid.
(9) Solution of a liquid in a liquid.

Of these types the solutions of gases in gases, gases in liquids, liquids in liquids, and solids in liquids are especially important.

Solutions of Gases in Gases. Every mixture of two or more gases can be considered a gaseous solution. As may be expected from the kinetic theory of gases, all nonreactive gases are miscible in all proportions, yielding solutions whose physical properties are very nearly additive provided the total pressure is not too high. Under the latter conditions the partial and total pressures are governed by Dalton's law, the partial and total volumes by Amagat's law. Both of these principles have already been discussed.[1]

---

[1] See pages 14–17.

Solubility of Liquids and Solids in Gases. The vaporization of a liquid and the sublimation of a solid into a gas phase may be considered as solution of these substances in a gas. These processes involve first the conversion of the liquid or solid to vapor, and the subsequent solution of the vapor in the gas. Because the vaporization and sublimation pressures of a substance are fixed at any given temperature, the amounts of liquid and solid that can vaporize into a given volume of gas are limited. Consequently, *at any given temperature* the solubility of liquids and solids in gases is limited to the amount necessary to establish the equilibrium pressures.

Solubility of Gases and Liquids in Solids. Gases and liquids may dissolve in solids to form apparently true homogeneous solutions. Examples are the solubility of hydrogen in palladium, and the solubility of liquid benzene in solid iodine. Both solutions formed are solid.

Solubility of Solids in Solids. True solutions composed of solid constituents are encountered quite frequently in chemistry and metallurgy. When two solids dissolve in each other, the solutions formed may be completely or partially miscible, depending on the nature of the substances involved and the temperature. Examples of salts forming solid solutions are potassium and ammonium sulfates, copper and ferrous sulfates, and the alums of ammonium and potassium. Many metal pairs likewise form solid solutions, as for instance gold and platinum, gold and palladium, silver and palladium, and copper and nickel. Since the formation of a solid solution would occur extremely slowly while both materials remained solid, it is necessary to resort to crystallization, either from solution in the case of salts, or from the molten materials in the case of metals, in order to obtain these solid solutions.

Temperature has no influence on the solubility when both substances involved in the solid solution formation are completely miscible in the solid state. When the two are only partially miscible, however, the extent of solubility depends on the temperature. Some substances may be completely miscible in the solid state at higher temperatures and only partially miscible at lower ones, a transition occurring from one type to the other. There are also other possibilities, but these will be discussed in greater detail under the subject of Phase Rule in Chapter XIV.

Solutions of Gases in Liquids. Gases dissolve in liquids to form true solutions. The degree of solubility is in all cases limited, and depends on the nature of the gas, the nature of the solvent, the pressure, and the temperature. Gases like nitrogen, hydrogen, oxygen, and helium dissolve in water only to a slight extent, while gases like hydrogen chloride and ammonia are very soluble. The large solubility in the latter cases is accounted for by the chemical reaction of these gases with the solvent to form hydrochloric acid and ammonium hydroxide, respectively. In

the first-mentioned gases there is no chemical interaction between solute and solvent, solution being due rather to the effect of attractive molecular forces of the solvent for the molecules of solute.

Solubility depends greatly also on the nature of the solvent. Gases like nitrogen, oxygen, and carbon dioxide are much more soluble in ethyl alcohol than they are in water at the same pressure and temperature, while hydrogen sulfide and ammonia are more soluble in water than in ethyl alcohol. Frequently chemical similarity between solute and solvent leads to a higher solubility, as is evidenced by the fact that hydrocarbon vapors dissolve more readily in hydrocarbon and other organic solvents than they do in water. Still, chemical similarity or dissimilarity is not an infallible criterion of solubility. Thus acetylene, which is quite different in chemical characteristics from water, dissolves to a greater extent in water at 0° C than does oxygen. When the nature of the solvent is modified by dissolving in it nonvolatile substances, especially electrolytes, the solubility of gases is generally decreased. The phenomenon, referred to as "salting out," is believed to be due to the fact that the amount of solvent available for solution is decreased by the removal of some of it through solvation of the dissolved substance.
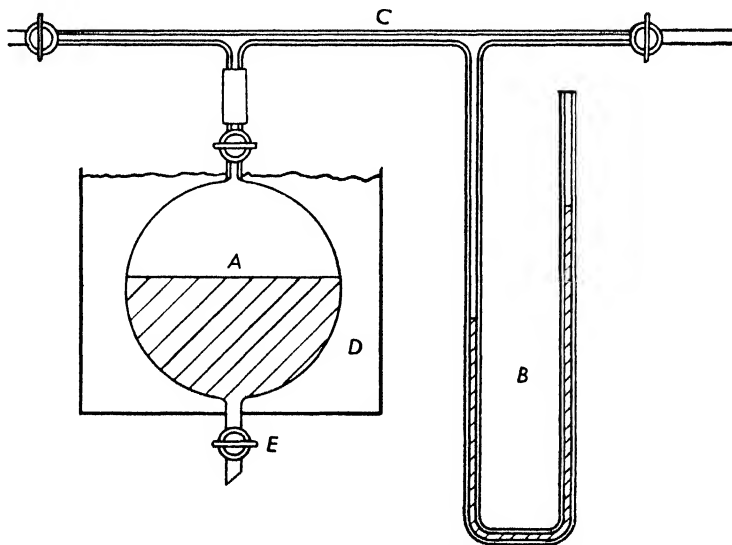
Fig. 1. Solubility of Gases by Static Method

Before discussing the effects of pressure and temperature, two simple methods for the experimental determination of solubility of gases in liquids will be described. For the determination of the solubility of sparingly soluble gases, particularly those which are difficult to estimate quantitatively with accuracy, the apparatus shown in Fig. 1 is especially

useful. The absorption bulb A of known volume V is attached to the capillary bridge C by a short piece of flexible tubing. Also connected to this bridge are the mercury manometer B and the gas reservoir (not shown). The bulb A is suspended in thermostat D. The system is first completely evacuated, and then gas,whose solubility is to be determined, is admitted until the desired pressure P is reached. Next a measured volume of gas-free solvent V_s is admitted into the bulb through stopcock E. The bulb is now agitated at intervals until equilibrium between the solution and gas phase, as indicated by the constancy of the manometer levels, is established. The final pressure P_f is then recorded.

To calculate from the above data the solubility of the gas in grams per unit volume of solvent we proceed as follows. If the gas is considered ideal, the number of moles of gas n introduced into A is

$$n = \frac{PV}{RT} \tag{2a}$$

where T is the temperature of the thermostat. The number of moles of gas, n_1, remaining undissolved at equilibrium is given by

$$n_1 = \frac{P_f(V - V_s)}{RT} \tag{2b}$$

Consequently, the number of moles of gas which dissolved in V_s cc of solvent is $(n - n_1)$, and hence the solubility S of the gas in grams per cubic centimeter at temperature T and pressure P_f is

$$S = \frac{M(n - n_1)}{V_s} \tag{2c}$$

where M is the molecular weight of the gas. If the solvent has a measurable vapor pressure, its value must be subtracted from the final pressure P_f in order to obtain the partial pressure of the test gas, and the corrected pressure used in equation (2b).

For highly soluble gases that can be analyzed readily in solution, such as hydrogen chloride and ammonia, the dynamic method originally proposed by Roscoe and Bunsen is very suitable. The apparatus is shown in Fig. 2. The saturation bulb A is partially filled with the solvent, and the apparatus is then suspended in a constant temperature bath B. The dried and purified gas to be dissolved is admitted into A through sidearm C, and the undissolved gas exits through arm D. When the solvent is saturated with gas, a sample for analysis is withdrawn slowly through the capillary stopcock E into the weighed flask F containing a measured quantity of standard reagent capable of reacting with the dissolved gas. The analysis is completed then by titration. This method has the

advantages over the preceding static method that manipulation is reduced to a minimum, and that several determinations of solubility can
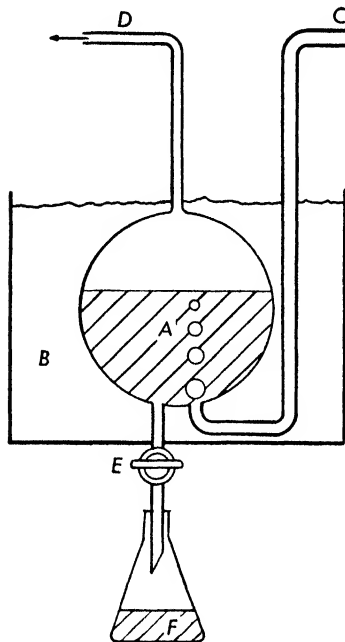
Fig. 2. Solubility of Gases by Dynamic Method

be made with a single filling without disrupting the flow of gas through the apparatus.

Henry's Law. The solubility of a gas in a liquid at any given temperature depends on the pressure of the gas and is, furthermore, directly proportional to it. The principle that *the solubility of a gas in a liquid is directly proportional to the pressure of the gas above the liquid at equilibrium* is known as *Henry's law*, and may be formulated as

$$C = kP \qquad (3)$$

where C is the concentration of the gas in solution, P the pressure of the gas above the solution at equilibrium, and k a proportionality constant known as *Henry's law constant*. From equation (3) it is evident that, when the pressure of a gas is doubled, the solubility is also doubled, etc. The magnitude of the constant k depends on the nature of the gas and solvent, the temperature, and the units in which C and P are expressed. It must be determined experimentally in each case.

When several gases are being dissolved simultaneously in a solvent, equation (3) is valid for each gas independently, provided C is the concentration and P the *partial pressure* of each gas. We may say, therefore, that *the solubility of each gas from a mixture of gases is directly proportional to the partial pressure of the gas in the mixture.* The proportionality constant k will, of course, be different for each gas.

The validity of Henry's law may be illustrated with the data given in Table 1 for the solubility of oxygen in water at several pressures at 25° C. If the law is correct the ratio $C/P = k$ should be constant at all pressures. This is actually the case, as may be seen from column 3 of the table.

The strict applicability of Henry's law, like that of all ideal laws, is limited to the lower pressures. At high pressures the law becomes less exact, and the proportionality constants exhibit considerable variation.

TABLE 1

SOLUBILITY OF OXYGEN IN WATER AT 25° C

P (cm Hg)	C (g/liter H_2O)	$\dfrac{C}{P} = k_{O_2}$
17.5	0.0095	0.000543
30.0	0.0160	0.000533
41.4	0.0220	0.000531
61.0	0.0325	0.000533
76.0	0.0408	0.000537

Generally, the higher the temperature and the lower the pressure, the more closely is the law obeyed. Furthermore, the law as given above is not applicable where the dissolved gas reacts with the solvent, or where the dissolved gas ionizes. When the ionization in solution is complete, the law breaks down altogether. The deviations in case of chemical reaction and partial dissociation can be readily understood and corrected, for when it is realized that Henry's law, which is a special case of a more general principle known as the Nernst distribution law, is valid only when it is applied to the concentration in solution of the *same molecular specie* as exists in the gas phase, and not to the total concentration in solution. Thus, when ammonia dissolves in water, part of the dissolved gas reacts to form ammonium hydroxide, which in turn dissociates partially to $NH_4^+ + OH^-$. The reactions involved may be written as

$$NH_3 \text{ (gas)} = NH_3 \text{ (dissolved)}$$
$$NH_3 \text{ (dissolved)} + H_2O = NH_4OH$$
$$NH_4OH = NH_4^+ + OH^-$$
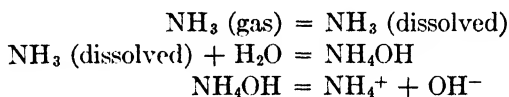
In the light of the above limitation, Henry's law to be applicable to the solubility of ammonia in water must be expressed not as $C_{total}/P_{NH_3} = k$, but as $C_{NH_3}/P_{NH_3} = k$, where C_{NH_3} is the concentration of ammonia in solution present as ammonia.

Effect of Temperature on Gas Solubility. The solubility of most gases in liquids decreases with increase in temperature, and consequently the Henry's law constants have smaller values at the higher temperatures. The extent of this decrease may be judged from Table 2, where the *solubility coefficients* for several gases in water are given at various temperatures. The solubility coefficient of a gas is defined as the number of cubic centimeters of gas, reduced to standard conditions, which dissolves in 1 cc of the solvent at a particular temperature and under a pressure of 1 atmosphere.

An interesting fact apparent from this table is that oxygen is about twice as soluble as nitrogen at any given temperature. At one time this difference in solubility was proposed as a possible method for preparing pure oxygen by successive solution of air in water. However, although

TABLE 2

SOLUBILITY COEFFICIENTS OF GASES IN WATER AT VARIOUS TEMPERATURES

Gas	0° C	10° C	25° C	50° C	100° C
CO_2	1.713	1.194	0.759	0.436	—
H_2	0.02148	0.01955	0.01754	0.01608	0.0160
N_2	0.02354	0.01861	0.01434	0.01088	0.0095
O_2	0.04758	0.03802	0.02831	0.02090	0.0170

a concentration of 97 per cent oxygen has been claimed after eight fractionations, the process is not sufficiently economical to be useful.

Because of the decrease in solubility at higher temperatures, liquids containing many types of dissolved gases may be purged of these by boiling. But this is not always the case. Experimental data show that some gases are more soluble at higher temperatures than at lower, and hence these are not readily removable by heating. In fact, dilute solutions of hydrogen chloride in water become more concentrated on boiling until eventually a solution containing about 20 per cent hydrogen chloride is attained.

Solutions of Solids in Liquids. The extent to which solids dissolve in liquids varies greatly with the nature of the solid and liquid, the temperature, and to a much lesser degree the pressure on the system. In all cases the limit of solubility is the concentration of the *saturated* solution. For any particular solute and solvent the concentration of the *saturated* solution at any given temperature and pressure is constant and does not depend on the manner in which the solution is prepared.

The concentrations of various solutes in a solvent necessary for saturation range over wide limits. Thus, at 20° C 100 grams of water dissolve 192 grams of ammonium nitrate, 222 grams of silver nitrate, 6.5 grams of mercuric chloride, and only 8.4×10^{-6} grams of silver bromide. In ethyl alcohol, on the other hand, the order of solubility of mercuric chloride and ammonium nitrate is reversed, 100 grams of the solvent dissolving 47.6 grams of mercuric chloride and only 3.8 grams of ammonium nitrate. As a rule most inorganic substances are more soluble in water than in organic solvents, while the reverse holds true for organic substances. There are, however, many exceptions.

The influence of temperature on the solubility of a solute in a particular solvent is, in general, quite pronounced, as may be seen from Table 3. Because most substances absorb heat on solution they tend to become more soluble at higher temperatures. When the solution process is exothermic, however, a decrease of solubility with temperature may

TABLE 3

SOLUBILITY OF SOLIDS IN WATER AT VARIOUS TEMPERATURES
(g/100 g H_2O)

Solid	0° C	20° C	40° C	60° C	100° C
NH_4Cl	29.4	37.2	45.8	55.2	77.3
$CaSO_4 \cdot 2\ H_2O$	0.176	—	0.210	0.205	0.162
$CuSO_4 \cdot 5\ H_2O$	14.3	20.7	28.5	40.0	75.4
$MgSO_4 \cdot 6\ H_2O$	40.8	44.5	—	53.5	74.0
KCl	27.6	34.0	40.0	45.5	56.7
KNO_3	13.3	31.6	63.9	110.0	246.
$AgNO_3$	122.	222.	376.	525.	952.
$NaCl$	35.7	36.0	36.6	37.3	39.8
Na_2SO_4	—	—	48 8	45.3	42.5

be expected — as is the case with sodium sulfate. Occasionally substances are encountered, such as $CaSO_4 \cdot 2\ H_2O$, whose solubility behavior is not quite so regular. The solubility of $CaSO_4 \cdot 2\ H_2O$ increases with temperature up to 40° C, passes through a maximum at this temperature, and then decreases. From this behavior we may expect the solubility of $CaSO_4 \cdot 2\ H_2O$ in water to be endothermic up to 40° C, and exothermic above this temperature. At 40° C, the temperature of maximum solubility, heat should be neither evolved nor absorbed, since at this temperature there is a transition from a process in which heat is absorbed to one in which heat is evolved.

When the solubility of any substance is plotted against temperature, the curve obtained is continuous as long as there is no change in the nature of the saturating solid phase. As soon as the solid phase changes, however, a break in the solubility curve appears, and a new solubility curve, originating at the point of break, is obtained, which gives now the solubility of the solid phase formed as a function of the temperature. For substances exhibiting such changes in solid phase the temperatures at which the breaks in the solubility curve occur are definite and characteristic of the substances involved; they represent the temperatures at which the original solid phase and the new solid phase are in equilibrium with the same solution. In other words, such a break occurs at the temperature at which the solution is saturated with respect to both solid phases.

The particular change in the nature of the solid phase involved may be a transformation of one crystalline form to another, a change from a hydrate to the anhydrous salt, or a transformation of one hydrate to another. These changes will be discussed more fully in Chapter XIV. At present only one example will be given, namely, the transformation

of β-rhombic ammonium nitrate to the γ-rhombic solid modification, which leads to a discontinuity in the water solubility curve at 32° C. Each of these forms of ammonium nitrate has its own solubility curve, but each form is stable over a different temperature interval. Below 32° C the β-form is stable and is the saturating phase, while above 32° C the γ-modification is the stable form. At 32° C, the transition temperature, both forms are stable. Hence the solubility curves of the two intersect, and the solution is saturated with respect to both. Any break in a solubility curve may be viewed, therefore, as the intersection of two distinct solubility curves, each corresponding to and characteristic of its particular saturating phase.

The effect of pressure on the solubility of solids in liquids is generally quite small. A change of 500 atmospheres in pressure increases the solubility of sodium chloride in water by only 2.3 per cent, and decreases the solubility of ammonium chloride by only 5.1 per cent. Another effect on the extent of solubility of a solid in a liquid which should be mentioned is that of the particle size of the solid. It has been observed that the solubility of a solid in a liquid is increased when the particle size of the saturating phase becomes very small. Thus, when the particle size of calcium sulfate is decreased from 2 to 0.3 micron the solubility in water at 25° C goes up from 2.085 to 2.476 grams per liter. Again, the solubility of barium sulfate is increased from 2.29×10^{-3} grams per liter of aqueous solution at the same temperature to 4.15×10^{-3} grams per liter when the particle size is reduced from 1.8 to 0.1 micron. These facts explain why it is necessary in analytical procedures to digest a precipitate in order to increase the particle size and thereby decrease the solubility. Any solution saturated with respect to fine particles will be supersaturated with respect to any coarser ones, and hence the tendency will be for solute to precipitate onto the coarser particles. The result will be that the crystal size of the larger particles will increase, while the finer particles will disappear by solution.

Solubility of Liquids in Liquids. Two liquids on being brought into contact with each other may be (1) completely miscible, (2) partially miscible, or (3) completely immiscible. As the classifications imply, completely miscible liquids dissolve in each other in all proportions, partially miscible liquids only in limited proportions, and immiscible liquids not at all. Strictly speaking, there probably is no pair of entirely immiscible liquids. Even mercury, when thoroughly dried over long periods of time to remove all traces of water, exhibits remarkable differences in some of its physical properties as compared to the ordinarily dried material. The distinction between partial miscibility and complete immiscibility is one of degree, and depends on the sensitivity of the analytical methods available for the detection of the substances involved.

The degree of miscibility of two liquids depends on their nature and the temperature. In general, two liquids will be the more miscible the closer their chemical similarity. Thus liquid pairs like benzene and toluene, water and ethyl or methyl alcohol, and water and acetone are completely miscible in each other, while water and nitrobenzene or water and chlorbenzene are not. Increased solubility or even complete miscibility sometimes can be attained by increasing the temperature, as with the liquid pair phenol and water. This, however, is not generally true. Triethylamine and water are completely miscible below 18.5° C, and only partially miscible above this temperature. There are also other types of partial miscibility, but these will be discussed in greater detail later.

The properties of binary liquid solutions in which we are particularly interested here are the vapor pressures and the distillation behavior at a constant pressure. These can be understood best in terms of the concept of an *ideal solution* to which all real solutions can be compared.

Ideal Solutions. Just as it was possible to set up an ideal gas as a criterion and to predict its properties from theoretical considerations, so, too, is it possible to define an ideal solution whose properties can be predicted on the basis of certain simple laws. An ideal solution may be defined as a solution in which the various pure constituents involved do not experience any modification of properties beyond that of dilution. In the formation of an ideal solution no heat will be evolved or absorbed on solution, and no heat effects will accompany the dilution of the solution with either constituent. For such solutions, also, the total volume will be equal to the sum of the volumes of the individual constituents, while other physical properties will be the average, depending on amounts present, of the properties of the pure constituents. And further, ideal solutions must obey strictly *Raoult's law of vapor pressures*, to be discussed in the next section, throughout the complete range of concentrations. As we shall see, some solutions approximate quite closely to these requirements; others, however, meet them only in dilute solutions.

Raoult's Law. Raoult's law deals at any given temperature with the partial vapor pressures above a solution of the constituents in solution. The law states that the *partial vapor pressure of any volatile constituent of a solution is equal to the vapor pressure of the pure constituent multiplied by the mol fraction of that constituent in solution.* Thus, if we have a solution of two binary, volatile, and miscible liquids of composition N_A and N_B, where N_A and N_B are the mol fractions of the two constituents, then the vapor pressures of the two constituents above the solution are, respectively,

$$P_A = P_A^0 N_A \tag{4a}$$
$$P_B = P_B^0 N_B \tag{4b}$$

where P_A^0 and P_B^0 are the vapor pressures of the *pure* substances A and B. Equations (4) are a mathematical statement of Raoult's law. From these equations the total vapor pressure above a solution such as described above follows as

$$P = P_A + P_B$$
$$= P_A^0 N_A + P_B^0 N_B \qquad (5)$$

Since, however, $N_A + N_B = 1$, then $N_A = 1 - N_B$, and

$$P = P_A^0(1 - N_B) + P_B^0 N_B$$
$$= P_A^0 - P_A^0 N_B + P_B^0 N_B$$
$$= (P_B^0 - P_A^0)N_B + P_A^0 \qquad (6)$$

Equation (6) gives the total vapor pressure above a solution as a function of the vapor pressures of the two pure substances and the mol fraction of one of these, N_B. This equation is of the form $P = mx + b$, where $m = (P_B^0 - P_A^0)$ and $b = P_A^0$, and hence it predicts that, when the total pressure of the solution is plotted against the mol fraction N_B, the curve should be a straight line with the indicated slope, and an inter-

cept at $N_B = 0$ of $b = P_A^0$. To show this graphically, it is convenient to use for binary mixtures a plot such as shown in Fig. 3. The ordinate is pressure, while the abscissa is N_B. Since N_B can have only values between zero and one, these must be the limits of the diagram. Further, since at $N_B = 0$ only pure A is present, and at $N_B = 1$ only pure B, the ordinate at $N_B = 0$ represents the relations for pure A, while the ordinate at $N_B = 1$ does the same for pure B. The ordi-

Fig. 3. Vapor Pressure of Miscible Liquids According to Raoult's Law

nates between these two limits represent the properties of all possible proportions of A and B, and hence a plot such as is shown in Fig. 3 gives the complete vapor pressure relations of all possible mixtures of the two substances.

In line with these requirements, the ordinates in Fig. 3 at $N_B = 0$ and $N_B = 1$ give P_A^0 and P_B^0, respectively. Since the pressures of all intermediate concentrations lie on a straight line that includes these points, the solid line in the diagram must represent the total vapor pressures of all possible mixtures of A and B. We see, therefore, that for binary solutions obeying Raoult's law the vapor pressures are intermediate between those of the pure components and lie on a straight line between these points.

The dotted lines in the diagram represent the *partial* pressures of the individual components and are, similarly, the plots of equations (4a) and (4b). These vary linearly from $P_A = P_A^0$ and $P_B = 0$ at $N_B = 0$, to $P_A = 0$ and $P_B = P_B^0$ at $N_B = 1$. At all intermediate mol fractions the total pressure ordinate is the sum of the partial pressure ordinates. Thus, at $N_B = x$, P_B is equal to the distance $a'x$, $P_A = b'x$, and the total pressure $P = c'x = a'x + b'x$.

It should be emphasized that the above relationships demanded by Raoult's law apply to the total and partial vapor pressures as a function of the mol fractions of the constituents *in solution*. To obtain the relation between the composition of a solution and the composition of the vapor above it, let Y_B be the mol fraction of B *in the vapor* above a solution of composition N_B. Then, according to Dalton's law of partial pressures,

$$P_B = Y_B P$$

$$Y_B = \frac{P_B}{P} \tag{7}$$

But $P_B = P_B^0 N_B$, while P is given by equation (6). Consequently,

$$Y_B = \frac{P_B^0 N_B}{(P_B^0 - P_A^0)N_B + P_A^0} \tag{8}$$

Equation (8) gives the mol fraction of B *in the vapor* above a solution of composition N_B in terms of the vapor pressures of the pure substances and N_B. It is readily seen that there is a definite composition of vapor corresponding to each composition of solution, but that the two are not the same and cannot be the same unless $P_A^0 = P_B^0$.

With the aid of equation (8) a vapor pressure-vapor composition curve can be constructed for solutions obeying Raoult's law which will show the composition of the vapor corresponding to any particular composition of solution. Such a curve calculated at a constant temperature for $P_A^0 = 147$ mm and $P_B^0 = 396$ mm Hg is shown in Fig. 4. The straight line gives the total pressure above the solution as a function of the mol fraction of B in solution, N_B, while the curve lying
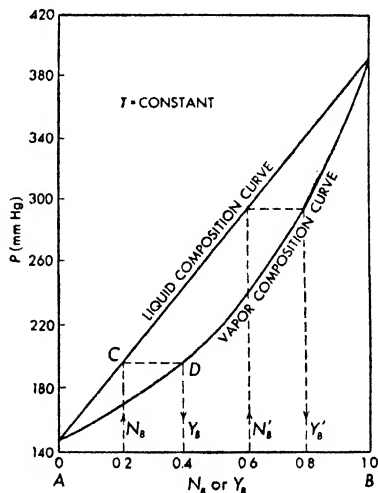
Fig. 4. Liquid and Vapor Composition Curves for an Ideal Solution

below it represents the total vapor pressure as a function of the mol fraction of B in the vapor, Y_B. To obtain the composition of vapor corresponding to, say, a solution for which $N_B = 0.2$, we move vertically to point C to obtain the total vapor pressure of the solution. This same pressure is given also on the vapor composition curve by point D, corresponding to a composition $Y_B = 0.402$, and hence for the liquids in question when $N_B = 0.2$, $Y_B = 0.402$. Similarly, when $N_B' = 0.6$, $Y_B' = 0.803$. It will be observed that the vapor always is richer in B than the solution, i.e., the vapor is richer in the more volatile component.

Vapor Pressure of Actual Liquid Pairs. A few binary miscible liquid systems obey Raoult's law strictly throughout the complete range of concentrations. One of these, the pair ethylene dibromide-propylene dibromide at 85.05° C, is shown in Fig. 5. Others are the pairs benzene-ethylene dichloride, carbon tetrachloride-stannic chloride, and chlorbenzene-brombenzene. Most systems, however, deviate from Raoult's law to a greater or lesser degree dependent on the nature of the liquids and the temperature.
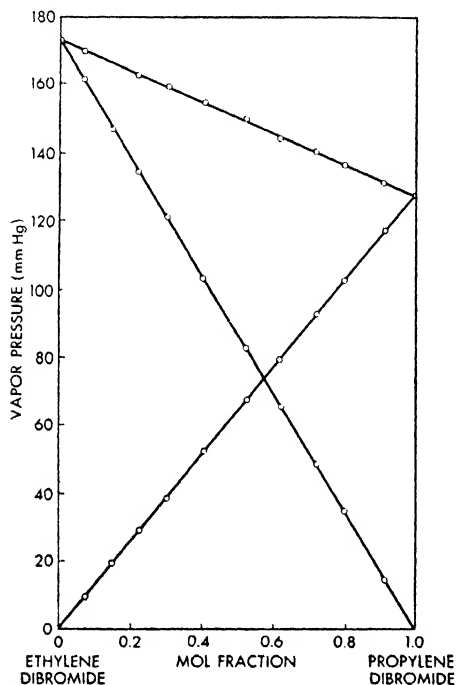
Fig. 5. Vapor Pressures of the System Ethylene Dibromide-Propylene Dibromide at 85.05° C

The character of the deviations from Raoult's law may be judged from Figs. 6, 7, and 8, which show the total and partial pressures in the several systems in their dependence on the mol fraction. In the pair carbon tetrachloride-cyclohexane, shown in Fig. 6, the total pressure as well as the partial pressures exhibit positive deviations from Raoult's law, but the total pressure is at all times intermediate between the vapor pressures of the two pure components. Similarly, in the system carbon disulfide-methylal, Fig. 7, the deviations from Raoult's law are positive, but the total vapor pressure curve rises to a *maximum* which is above the vapor pressure of either pure constituent. On the other hand, the pair chloroform-acetone, Fig. 8, exhibits negative deviations from Raoult's law which lead to a *minimum* in the total vapor pressure of the system,
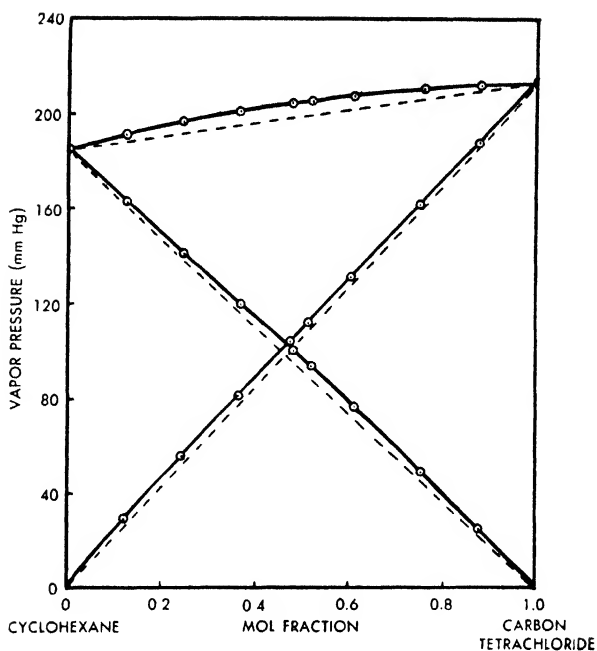
Fig. 6. Vapor Pressures of the System Cyclohexane-Carbon Tetrachloride at 40° C
[Scatchard, Wood, and Mochels — J. Am. Chem. Soc., 61, 3208 (1939)]
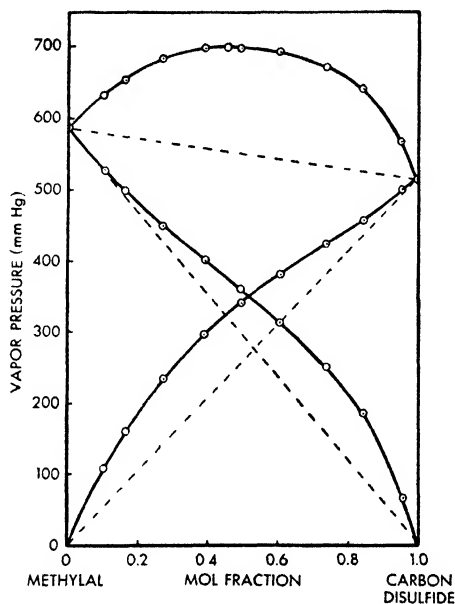
Fig. 7. Vapor Pressures of the System Methylal-Carbon Disulfide at 35.2° C

i.e., the vapor pressures of certain concentrations of the solution are *below* the vapor pressures of either of the pure constituents.

The vapor pressure curves given in Figs. 6, 7, and 8 are typical of the vapor pressure behavior of volatile liquid pairs. On the basis of these diagrams, all miscible liquid pairs can be classified into three general types, namely:
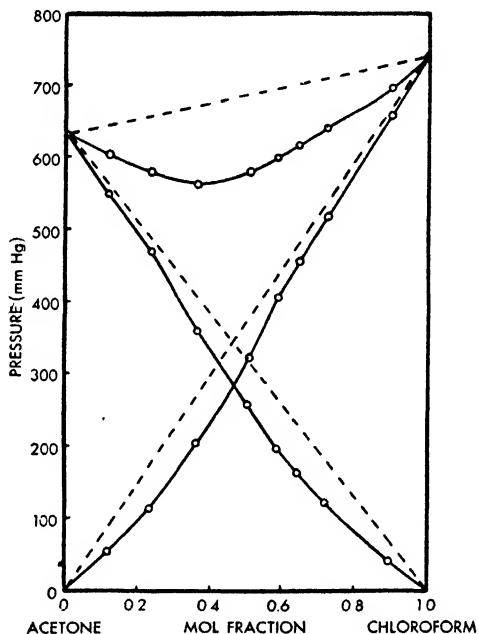
Fig. 8. Vapor Pressures of the System Acetone-Chloroform at 55.1° C

Type I: Systems whose total vapor pressure is *intermediate* between those of the pure components. Examples: Carbon tetrachloride-cyclohexane, carbon tetrachloride-benzene, benzene-toluene, water-methyl alcohol.

Type II: Systems exhibiting a *maximum* in the total vapor pressure curve. Examples: Carbon disulfide-methylal, carbon disulfide-acetone, benzene-cyclohexane, benzene-ethyl alcohol, chloroform-ethyl alcohol, and water and ethyl or *n*-propyl alcohol.

Type III: Systems exhibiting a *minimum* in the total vapor pressure curve. Examples: Chloroform-acetone, methyl ether-hydrogen chloride, pyridine-acetic acid, and water and formic, nitric, hydrochloric, or hydrobromic acids.

The vapor pressure-*liquid* composition curves of these various types along with the corresponding vapor pressure-*vapor* composition curves are illustrated in Fig. 9. All these curves correspond to a constant temperature.

It will be observed that for all types the vapor composition curves lie *below* the liquid composition curves, and that in types II and III the two

curves are in contact at the points of maximum and minimum vapor pressure respectively. At these points, therefore, the compositions of the solution and vapor are identical, and the diagrams of types II and III may be thought of as being composed of two curves of type I, one for pure A and solution of maximum or minimum vapor pressure, C or D, the other of solution of maximum or minimum vapor pressure, C or D,
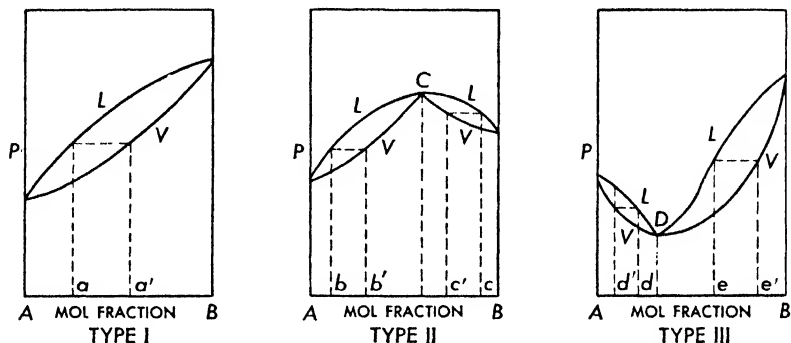
Fig. 9. Types of Binary Vapor Pressure Diagrams (Temperature Constant)

and pure B. The position of the vapor composition curves with respect to the liquid composition curves follows from the fact that out of a mixture of two volatile substances in solution the one of higher vapor pressure will volatilize to a greater extent than the one of lower vapor pressure, yielding a vapor of composition richer in the more volatile constituent than is the solution. Thus, in the liquid pair of type I shown in Fig. 9, B is more volatile than A, and consequently the vapor above all concentrations of solution must be richer in B than is the solution. Hence, the composition of vapor corresponding to point a on the diagram must lie closer to B than point a, and this is possible only if the vapor composition curve is below the liquid composition curve. As the diagram indicates, the composition of vapor corresponding to a is a', a mixture considerably richer in B than is a. For the same reason all mixtures between A and C in type II must have vapors richer in B than the solution, while all mixtures between C and B must be richer in A than the solution. These conditions are satisfied by the diagram as drawn, for the vapor composition corresponding to a point such as b is b', richer in B than b, while that of composition c is c', richer in A than c. Similarly the compositions of vapor in type III must be richer in A than the solutions between A and D, while richer in B between D and B. These conditions are again satisfied by the diagram drawn, as may be verified by the two points indicated, d with vapor composition d', and e with vapor composition e'.

Determination of Vapor Pressure-Composition Diagrams. A word may be in order on the experimental methods employed to deter-

mine vapor pressure-composition diagrams of binary solutions such as those described. In general these methods are quite similar to those employed to determine the vapor pressure of pure liquids at a given temperature, with the sole exception that both the liquid and gaseous phases in equilibrium must be analyzed to determine their compositions. Once the composition of the vapor and the total vapor pressure corresponding to a liquid composition are known, the partial vapor pressures of the constituents above the solution can be calculated with the aid of Dalton's law of partial pressures. By repeating the measurements of total vapor pressure and the analyses for a number of liquid compositions ranging from pure A to pure B, sufficient data may be accumulated to permit the construction of a complete vapor pressure-composition diagram at a given temperature. Since the vapor pressures of both the pure constituents and the various possible concentrations of solution vary with temperature, a number of such diagrams at various temperatures must be constructed before the complete vapor pressure behavior of the solutions at various temperatures can be defined.

Boiling Point Diagrams of Miscible Binary Mixtures. Because at any given temperature the vapor above any solution is richer in the more volatile substance than the solution, a solution can be made to shift in composition toward the less volatile constituent by removing the vapor above it. Again, if the vapors are condensed, and the new vapor above the condensate removed, the new vapors will be found considerably richer in the more volatile constituent than are the solutions from which they came. By repeating this process it is possible to obtain a concentration of the more volatile constituent in the vapor and a concentration of the less volatile constituent in the solution. Such a process of concentrating the constituents is known as fractional distillation; and, since the process described takes place at constant temperature, it may be designated as an isothermal fractional distillation.

In practice it is much more convenient to conduct a distillation at *constant pressure* rather than at constant temperature; i.e., it is much easier to conduct a distillation such that the temperature varies while the pressure remains constant than it is to hold the temperature constant and let the pressure vary. At a given confining pressure any solution of definite composition will boil at a temperature at which its *total* vapor pressure becomes equal to the confining pressure. If we designate by P the confining pressure, then the condition for boiling may be written as

$$P_A + P_B = P_T = P \qquad (9)$$

Thus, at atmospheric pressure a solution will boil at the temperature at which its total vapor pressure becomes equal to 760 mm Hg. Since different compositions of a solution have different vapor pressures, it must

follow that the various solutions will not reach a total vapor pressure
equal to the confining pressure at the same temperature, and therefore
solutions of various concentrations will boil at different temperatures.
In general, solutions of *low* vapor pressure will boil at temperatures
higher than solutions whose vapor pressure is high, for solutions of high
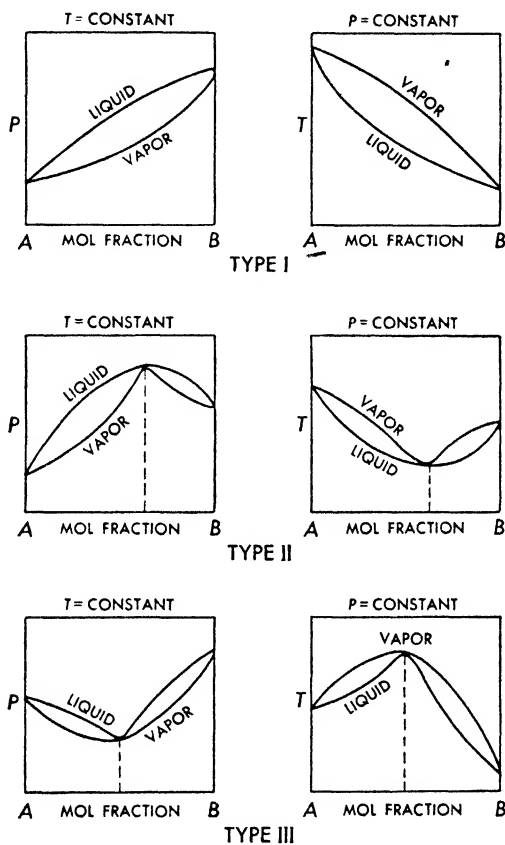
Fig. 10. Types of Distillation Diagrams Corresponding to Various Vapor Pressure
Diagrams

vapor pressure can reach a total pressure equal to the confining pressure
at relatively lower temperatures than solutions whose vapor pressure is
low.

This latter fact permits the construction of the various types of tem-
perature-composition diagrams which will correspond to the three gen-
eral types of vapor pressure-composition diagrams already discussed.
These are shown in Fig. 10. In type I the vapor pressure of A is the
lowest pressure in the system, that of B is the highest, while the vapor
pressure of all possible compositions of A and B are intermediate between

the two. Consequently, the boiling point at constant pressure of A will be the highest in the system and that of B the lowest, while those of all compositions of A and B will be intermediate, and will be given by the liquid composition curve in the figure. Since the vapor coming off from any particular composition of solution must be richer in the more volatile constituent B, the vapor composition at any temperature must lie *closer* to B than the corresponding liquid composition, and hence the vapor composition curve must lie now *above* the liquid composition curve, as shown. The same considerations apply to the other two types. In type II the vapor pressure of the system is a maximum for composition C, and hence such a solution will boil at the lowest temperature, leading to a minimum in the boiling point curve. Again, since in type III the solution of composition D has the lowest vapor pressure in the system, it will boil at the highest temperature, and consequently the boiling point curve exhibits a *maximum*. In all cases the vapor composition curves lie above the liquid composition curves for the reasons given. We see, therefore, that any system whose vapor pressures are intermediate between those of the pure constituents will have a distillation diagram with intermediate boiling points, such as type I. On the other hand, any system of the maximum vapor pressure type will give a distillation diagram with minimum boiling point, while any system of the minimum vapor pressure type will have a distillation diagram with a maximum in the boiling point curve.

Distillation of Binary Miscible Solutions. Because of the differences in the distillation diagrams of the three types of solutions, the behavior of these on constant pressure distillation will be different. Consider first the behavior of a system of type I (Fig. 11). If we heat a solution of composition a, no boiling will start until temperature T_a is reached. At this temperature the vapor coming off from a and in equilibrium with it

Fig. 11. Distillation Behavior of Solutions of Type I

will have the composition a'. Since a' is richer in B than a, the composition of the residue must become richer in A, say b. The new composition of residue, b, cannot boil, however, until temperature T_b is reached, which is higher than T_a. In turn the vapor coming off from b will have the composition b', again richer in B, and consequently the composition of the residue will again be enriched in A, and again

the temperature must rise before the residue will boil. We see, therefore, that if the process described is continued the boiling point of the solution will rise from the initial boiling point T_a toward the boiling point, T_A, of pure A. At the same time the composition of the residue becomes richer in A than the original solution, and if the process is continued sufficiently long, a final residue of pure A can be obtained.

Suppose we consider now what can be done with the vapors. If the initial vapors obtained from the solution, namely, a', are condensed and again distilled, the boiling point of the new solution will be T'_a, and the composition of the distillate will be given by a''. This distillate is again richer in B than the original. If the process of condensing and redistilling is continued, eventually a vapor can be obtained composed essentially of pure B. Therefore, on distillation of any mixture of type I it is possible to separate eventually the constituents into a residue of the less volatile constituent A, and a distillate of the more volatile component B; i.e., *the two constituents forming a solution of type I can be separated by fractional distillation into the pure components.*

Such a separation into the pure components is impossible, however, with solutions of either type II or type III. Consider the distillation behavior of a solution of type II, Fig. 12, of which the system water (A) -ethyl alcohol (B) is an outstanding example. If a solution of composition between A and C, such as a, is distilled, the vapor coming off will have the composition a' and will be richer in B than the residue. Be-

Fig. 12. Distillation Behavior of Solutions of Type II

cause of this fact the composition of the residue will shift toward A, and hence the residue will have to boil at a temperature higher than the original solution a. If the distillation is continued, the same argument as was employed for solutions of type I indicates that eventually a residue of pure A, boiling at temperature T_A, will be obtained. On the other hand, if the vapors from the original solution, a', are condensed and redistilled repeatedly, a vapor of composition C will eventually be obtained. Such vapor when condensed and distilled will again yield vapor of composition C; i.e., the vapors coming off from the solution will have the same composition as the solution, and hence no further separation is possible by distillation. Consequently, *any mixture having a composition between A and C can be separated by fractional distillation only into a residue*
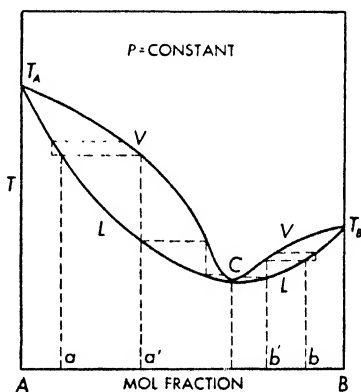
of pure A and a final distillate of composition C. No pure *B* can be recovered. Thus, in the system water-ethyl alcohol point *C* corresponds to a minimum boiling temperature of 78.13° C and a composition of 95.57 per cent ethyl alcohol by weight, and hence any solution having a composition between that of pure water and 95.57 per cent ethyl alcohol can be separated only into a residue of pure water and a constant minimum boiling mixture of 95.57 per cent alcohol in the distillate. No pure ethyl alcohol can be recovered.

On the other hand, if a solution of composition between *C* and *B* is distilled, for example, *b*, the vapor coming off, *b'*, will be *richer in A* than the original solution, and hence on repeated distillation the residue will tend toward pure *B*, while the distillate will tend toward *C*. Such solutions on complete distillation will yield, therefore, pure component *B* in the residue and constant boiling mixture *C* in the distillate. No *A* can be recovered by distillation. Thus alcohol-water mixtures richer in alcohol than 95.57 per cent can be separated by distillation only into pure alcohol in the residue and constant minimum boiling mixture *C*, 95.57 per cent alcohol, in the distillate. No amount of distillation will yield pure water.

Boiling solutions of composition *C* will yield, of course, vapor of exactly the same composition as the solution. Hence no separation of such solutions can be accomplished by the distillation of the binary mixture.

In summary it may be stated that all starting compositions of solutions belonging to type II will yield eventually a distillate of composition *C*. The residue will be either pure *A* or pure *B*, depending on whether the starting composition is between *A* and *C* or between *C* and *B*. A starting composition *C*, however, cannot be separated by distillation.

The behavior of solutions of type III on distillation will be analogous to that of solutions of type II, with the exception that *the residues tend toward the maximum boiling mixture, while the distillates tend toward the pure constituents* (Fig. 13). If the starting mixture has a composition between *A* and *D*, such as *a*, the vapor obtained on distillation, *a'*, will be richer in *A* than the solution. Hence the composition of the residue will shift toward *D* and will eventually reach it. A redistillation of the vapor, on the other hand, will finally yield a distillate of pure *A*. A mixture between *D* and *B*, such as *b*, however, will yield on distillation a vapor of composition *b'*, richer in *B* than the solution. Hence again the residue will shift toward *D*, while on redistillation the vapors will tend toward pure *B*. Consequently, complete distillation of a mixture such as *b* will eventually yield a residue of composition *D* and a distillate of pure *B*. We see, therefore, that any binary system of the maximum boil-

ing point type can be separated on complete fractional distillation into a residue of composition D, that of the constant maximum boiling mixture, and a distillate of either pure A or pure B, depending on whether the starting composition is between A and D or D and B. But a mixture of composition D cannot be separated further by distillation.

An example of a binary system exhibiting such a behavior is the pair water-hydrochloric acid. Although water boils at 100° C and hydrogen chloride at − 85° C under a pressure of one atmosphere, a solution of 20.24 per cent hydrochloric acid in water boils at 108.5° C. Any solution containing less than 20.24 per cent hydrochloric acid can

Fig. 13. Distillation Behavior of Solutions of Type III

be separated on distillation into a constant boiling mixture of 20.24 per cent hydrochloric acid and a final distillate of pure water, but no pure hydrogen chloride can be recovered from such solutions. On the other hand, any solution richer in hydrochloric acid than 20.24 per cent can be distilled to yield again the same constant composition residue, and, if the distillation is carried down to − 85° C, pure hydrogen chloride.

Azeotropic Mixtures. The constant boiling mixtures described above which exhibit either a minimum or a maximum boiling point are referred to either as *azeotropes* or *azeotropic mixtures.* At one time it was believed that such mixtures correspond to the formation of definite compounds between the two constituents. However, a definite compound should have a definite composition over a given temperature and pressure range. This is not the case with azeotropic mixtures. When the total pressure at which the solution boils is changed, it is found that not only the temperature at which the azeotrope is obtained, but also the *composition* changes, as may be seen from Table 4. Since this is not the behavior to be expected from a definite compound, the notion of such compound formation is at present discarded. It is believed, rather, that the particular behavior observed is the result of the interplay of intermolecular forces in solution.

The composition of azeotropic mixtures is remarkably constant at any given pressure. Advantage is taken of this constancy to prepare solutions of hydrochloric acid of very accurately known composition. A solution of hydrochloric acid of composition lower than that of the maximum boiling mixture is distilled until the azeotropic composition is reached. The azeotrope is then collected, and the barometric pressure
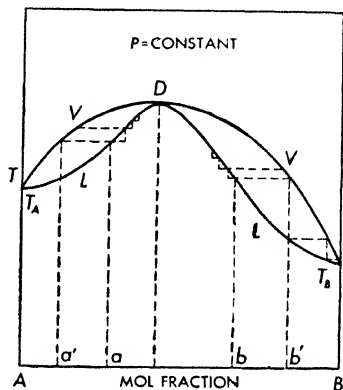
TABLE 4

EFFECT OF PRESSURE ON COMPOSITION OF AZEOTROPE IN SYSTEM H_2O-HCl

Pressure (mm Hg)	Per Cent by Weight of HCl in Constant Boiling Mixture
730	20.314
740	20.290
750	20.266
760	20.242
770	20.218

at which distillation took place is observed. From tables, such as Table 4, the composition of the azeotrope is then known, and any desired composition of hydrochloric acid can be prepared by dilution with an accuracy high enough for use in quantitative analysis without further standardization.

Table 5 lists the boiling points of various maximum and minimum constant boiling mixtures as well as the compositions of these, all for a total pressure of 760 mm Hg.

TABLE 5

BOILING POINTS AND COMPOSITIONS OF AZEOTROPIC MIXTURES
(pressure = 760 mm Hg)

Type	A	B	Boiling Point (°C)	Weight % of B in Azeotrope
Minimum B.P.	Water	Ethyl alcohol	78.15	95.57
	Water	n-Propyl alcohol	87.72	71.70
	Ethyl alcohol	Benzene	68.24	67.63
	Acetic acid	Benzene	80.05	98.
	Carbon disulfide	Ethyl acetate	46.1	3.
	Pyridine	Water	92.6	43.
Maximum B.P.	Water	Nitric acid	120.5	68.
	Water	Hydrochloric acid	108.5	20.24
	Water	Hydrobromic acid	126.0	47.5
	Water	Hydriodic acid	127.0	57.
	Water	Hydrofluoric acid	120.0	37.
	Water	Formic acid	107.1	77.
	Chloroform	Acetone	64.7	20.
	Pyridine	Formic acid	149.0	18.

The Fractionating Column. The type of distillation described heretofore, in which the vapor removed is in equilibrium with the total mass of boiling liquid, is designated as *equilibrium distillation*. The process of separating mixtures by distillation would be extremely complicated and tedious if it had to be performed by repeated distillations and condensations in a discontinuous manner. Instead, the separation is per-

formed in a continuous operation, known as *fractional distillation*, utilizing a distilling apparatus called a *fractionating column*, Fig. 14. The fractionating column consists essentially of three parts: a heated still A; the column proper D, composed of a series of plates whose detailed construction is shown in the figure; and a condenser F. The preheated mixture to be distilled is admitted through E onto one of these plates, and overflows through 2 to the plate below. On this lower plate the liquid comes in contact with vapor moving upward from the still through the "bubble caps" 3′ and 4′. These caps are so designed that the vapor must bubble through the layer of liquid on each plate before it can escape. In doing so, part of the less volatile constituent is condensed out of the vapor, and part of the more volatile constituent is vaporized out of the liquid. The vapor moving on to the next higher plate through 3 and 4 is richer, then, in the more volatile constituent than the vapor which approached the plate from below, while the liquid overflowing to the next lower plate through 1′ is richer

Fig. 14. Schematic Diagram of Fractionating Column

in the less volatile constituent than the liquid which reached the plate from above. The net result of the interaction between the vapor and the liquid at the plate is, therefore, a redistribution in favor of the more volatile constituent in the vapor and the less volatile constituent in the liquid; i.e., each plate acts essentially as a miniature still.

Since this process repeats itself at each plate, it is possible with a sufficient number of plates to separate the mixture into two end fractions, a residue of the less volatile component running into still A, where it can be drawn off, and a vapor passing from the top of the column containing essentially the more volatile constituent. This vapor is fed into a condenser F, where it is liquefied. Part of this liquid is drawn off through H, while part, the reflux, is returned to the column through G in order to maintain the stock of essentially pure distillate on the upper plates.

Ratio of Distillate to Residue on Distillation. For various chemical and chemical engineering calculations it is frequently necessary

to know the ratio of weight of distillate to weight of residue at each stage of an equilibrium distillation. For a binary mixture this information is readily available from the distillation diagram.
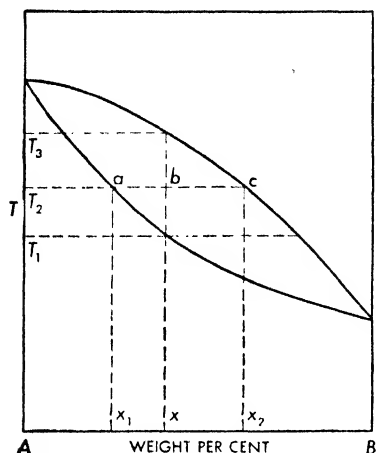
Consider a binary system whose distillation diagram is given by Fig. 15. The abscissa is expressed now as weight per cent rather than mol fraction. If we start with a mixture whose weight per cent is x, the system will be entirely liquid until temperature T_1 is reached. At T_1 the solution will begin to boil, and since the liquid becomes less concentrated in B, the boiling point will gradually rise as distillation proceeds. At a temperature such as T_2, if no vapor is removed from the system, the latter will still have the same overall composition as the starting mixture x, but it will be composed now of liquid of composition x_1 and vapor of composition x_2. Under such conditions the weight of liquid present is proportional to the linear distance $\overline{bc}$, while the weight of vapor present is proportional to the distance $\overline{ab}$, i.e.,

$$\frac{\text{Weight of liquid}}{\text{Weight of vapor}} = \frac{\overline{bc}}{\overline{ab}} \qquad (10)$$

This relation between weights and distances is perfectly general, and can be applied to any portion of the diagram at all temperatures. It must be remembered, however, that, when the abscissa is not weight per cent but mol fraction, the ratios of the distances represent the ratios of the number of *moles* of liquid and vapor rather than the weights.

Solubility of Partially Miscible Liquid Pairs

(a) *Maximum Solution Temperature Type.* When a small quantity of aniline is added to water at room temperature and the mixture is shaken, the aniline dissolves in the water to form a single liquid phase. However, when larger quantities of aniline are added, it is observed that instead of a single phase two liquid layers are formed. One of these, the lower, consists of a small amount of water dissolved in the aniline, while the upper consists of a small amount of aniline dissolved in water. Further addition of aniline to the system causes the upper, or water-rich, layer to diminish in size until finally it disappears, leaving only a single liquid phase composed of water in the aniline.

Fig. 15. Ratio of Distillate to Residue on Distillation

If this experiment is performed at constant temperature, it is found that, *as long as the two layers are present, the ratios of aniline to water in the two layers remain constant;* i.e., the compositions of the two layers, although different from each other, remain constant as long as two phases are present. The addition of small amounts of either aniline or water merely changes the relative volumes of the two layers, not their composition. As the temperature is raised, this behavior is found to persist except that the mutual solubility of the two liquids increases. When the temperature finally reaches 168° C, the compositions of the two layers become identical, and thereafter the two liquids are completely miscible. In other words, at 168° C and above aniline and water dissolve in each other in all proportions and yield only a single liquid layer on mixing.

This variation of the mutual solubility of water and aniline with temperature is illustrated in Fig. 16. At a temperature such as 100° C, point A represents the composition of the water-rich layer, while point A_1 represents the composition of the aniline-rich layer in equilibrium with A. Between A and A_1 all compositions yield two layers of compositions A and A_1. Outside these compositions the two liquids are mutually soluble at 100°; i.e., all compositions between pure water and A yield
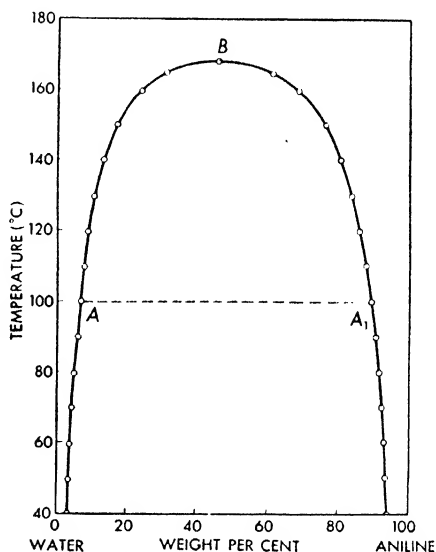
Fig. 16. Mutual Solubility of Water and Aniline at Various Temperatures

a solution of aniline in water, while all compositions between A_1 and pure aniline yield a solution of water in aniline. Since the same argument holds at other temperatures, it must follow that the dome-shaped area represents the range of existence of two liquid phases, the area outside the dome that of a single liquid layer. It should be noted that at 168° C the compositions of the two liquid layers become identical, both being given by B, and that thereafter the two liquids are completely miscible at all concentrations. The temperature corresponding to point B, i.e., the temperature at which solubility becomes complete, is called either the *critical solution temperature* or the *consolute temperature*.

(b) Minimum Solution Temperature Type. Figure 17 shows the effect of temperature on the mutual solubility of triethylamine and

water. The two liquids are completely miscible at or below 18.5° C, but only partially miscible above this temperature. Thus at 30° C, for instance, a solution of 5.6 per cent triethylamine in water is in equilibrium with one containing 4 per cent water in triethylamine. The temperature at which the two liquids become completely miscible is called in this case the minimum critical solution temperature, since the curve confining the area of partial miscibility exhibits a minimum.
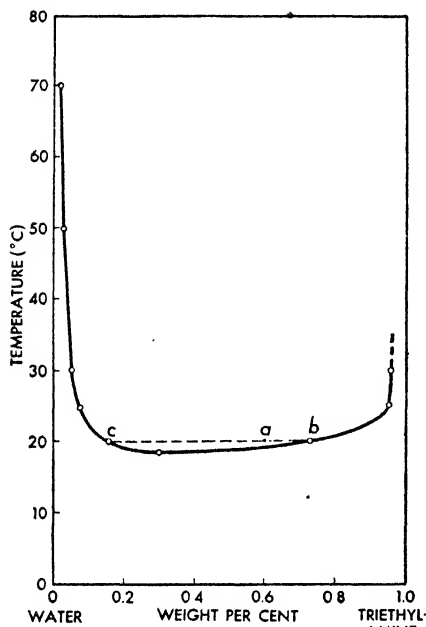
Fig. 17. Solubility of Triethylamine in Water at Various Temperatures

(c) *Maximum and Minimum Solution Temperature Type.* The system nicotine-water exhibits two critical solution temperatures, an upper and a lower, as may be seen from Fig. 18. Within the enclosed area the liquids are only partially miscible, while outside the enclosed area they are completely miscible. The upper or maximum solution temperature, point *C*, is 208° C, while the lower, or minimum solution temperature, point *C'*, is 60.8° C. The compositions corresponding to *C* and *C'* are the same, 34 per cent nicotine. At 94 to 95° C, point *A*, nicotine is least soluble in water, while water is least soluble in nicotine at 129 to 130° C, point *B*.

It has been found that on applying external pressure to this system the upper and lower critical solution temperatures approach each other, until a pressure is finally reached at which the two liquids become completely miscible.

(d) *Type without Critical Solution Temperature.* A final variation of these types is exhibited by the system ethyl ether-water, which has neither an upper nor a lower critical solution temperature. The two liquids are therefore only partially soluble in each other at the temperatures over which the solution exists.

It should be pointed out that it is always possible to deduce from solubility diagrams such as those given in Figs. 16, 17, and 18 the proportions by weight of the two layers present at equilibrium at the various temperatures. Consider specifically a system composed of 60 grams of triethyl-

amine and 40 grams of water present in equilibrium at 20° C. Since the percentage of amine by weight is 60 per cent, the overall composition is represented in Fig. 17 by point *a*. But this system is composed of solutions of compositions *b* and *c*, and hence, by the method used before,

$$\frac{Wt \text{ amine layer}}{Wt \text{ water layer}} = \frac{\text{distance } ca}{\text{distance } ab}$$
$$= \frac{60 - 15.5}{73 - 60}$$
$$= 3.42$$

Determination of Degree of Miscibility. Several methods are available for determining the mutual solubility of partially miscible liquids. These are: (*a*) the analytical method, (*b*) the Alexejew method, and (*c*) the titration method.

Fig. 18. Solubility of Nicotine in Water at Various Temperatures

(*a*) **The Analytical Method.** The two components are mixed and brought to equilibrium at some given temperature. Samples of each of the layers are then weighed out, and analyzed for either one or both constituents. The critical solution temperature is determined by successively changing the temperature and determining the composition until the two layers become completely miscible. This procedure is rather tedious, but yields accurate results when good analytical methods for the solution constituents are available.

(*b*) **The Alexejew Method.** In this method weighed quantities of both components are sealed in a glass tube, and the latter is immersed in a suitable bath. The temperature of the bath is then slowly raised until the two layers become completely miscible. It is essential to agitate the tube vigorously throughout the heating process. Since the overall composition of the mixture is known, the determination of the temperature at which the mixture becomes completely miscible yields directly a point on the temperature-composition curve. By repeating this procedure with various compositions of mixture, it is possible to trace out the solubility diagram without recourse to analysis.

(c) The Titration Method. In this method small measured amounts of constituent B are progressively added to a definite amount of A kept at a constant temperature. After each addition the flask is shaken vigorously, and the solution is then observed for the appearance of turbidity due to immiscibility. As soon as such turbidity is detected, the weight of B added is noted. From the weights of A and B thus obtained the composition at which immiscibility just occurs can be calculated. If the addition of B is continued until the solutions become again completely miscible, it is possible to determine the composition of the saturated solution of A in B as well. By repeating this procedure at various temperatures, the complete solubility curve can be determined.

For the estimation of mutual solubility it is very important that all components be very pure, as even small quantities of foreign substances may greatly influence the results.

Vapor Pressure and Distillation Diagrams of Partially Miscible Liquid Pairs. Although three types of vapor pressure and distillation diagrams are possible for partially miscible liquid pairs, only one of these is of sufficient importance to warrant discussion here.

A typical total vapor pressure-composition diagram for a pair of partially miscible liquids is shown in Fig. 19(a) specifically for the system *n*-butyl alcohol-water. Starting with pure butyl alcohol, whose vapor pressure is D', we observe an increase in the total pressure of the system on addition of water. This increase continues until the pressure given by point C', corresponding to a saturated solution of water in butyl alcohol, is reached. Similarly, starting with pure water of vapor pressure A', addition of butyl alcohol raises the vapor pressure until a saturated solution of butyl alcohol in water, point B', is reached. The vapor pressure corresponding to C' is *exactly the same* as that of B'; i.e., the saturated solution of butyl alcohol in water has exactly the same vapor pressure as a saturated solution of water in butyl alcohol. And, furthermore, the total vapor pressure above the system in the region of partial miscibility, between B' and C', where two layers are present, is *constant and equal to that of either of the two layers, and not the sum of the two.* This state of affairs is a direct consequence of the thermodynamic conditions of equilibrium in such a system. For equilibrium to exist in a system involving the distribution of a particular component between two phases, such as butyl alcohol between the upper and lower layers, the vapor pressure of the particular constituent above each of the layers must be the same, and the vapor pressure above both must be that of any one. Consequently, the vapor pressure of butyl alcohol above B' must be equal to that above C', and the total must be the same as *either* that of B' or C'. Since the same considerations apply to the water in the two phases, the total pressure above both layers is

the same as that of B' or C', and is constant as long as both phases are present.

The curves $A'H'$ and $D'H'$ give the compositions of vapor in equilibrium with $A'B'$ and $D'C'$, respectively. For all compositions of liquid between B' and C' the composition of vapor is constant and equal to H'.
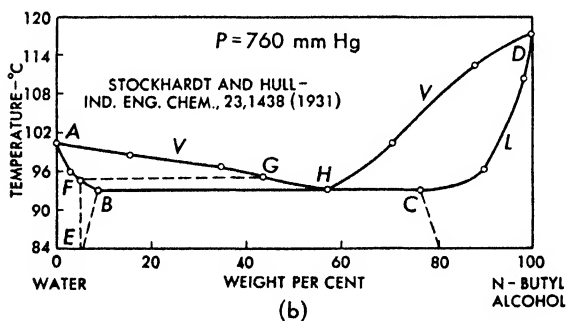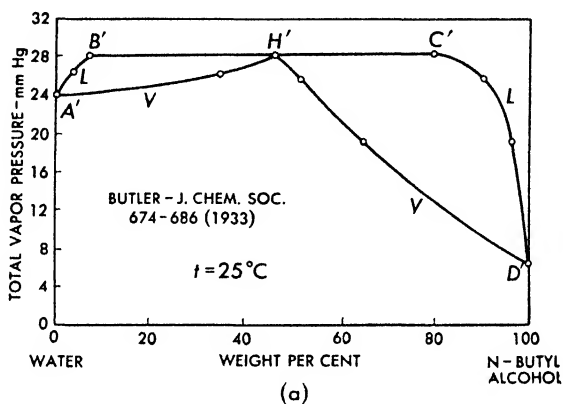
Fig. 19. Vapor Pressure and Distillation Diagrams for the System *n*-Butyl Alcohol-Water

The distillation diagram which corresponds to a system exhibiting the vapor pressure behavior shown by butyl alcohol-water is shown in Fig. 19(*b*). The curve CD gives the boiling points of all solutions of water in butyl alcohol, while HD gives the compositions of vapor corresponding to the various compositions of the liquid phase. Similarly AB gives the boiling points of all solutions of butyl alcohol in water, while AH gives the corresponding vapor compositions. The boiling points of all overall compositions that yield two layers, namely, between B and C, are given by BC, while the composition of vapor that corresponds to BC is given by point H; i.e., as long as the two saturated layers are pres-

ent, the boiling point of the system is *constant,* and the composition of the vapor coming off is also *constant and independent of the overall composition.*

To understand the distillation behavior of such a system, consider first the behavior of solutions between A and B in composition, such as E. This solution will boil when temperature F is reached, and the composition of vapor coming initially from the boiling solution will be G. Since the vapor is richer in alcohol than the solution, the composition of the liquid phase will shift toward that of water, and hence the boiling point will rise on distillation and will approach that of pure water. In an analogous manner all solutions between D and C will boil with evolution of vapor of composition along DH, and since this vapor is richer in water than the liquid, the boiling points will rise on distillation toward D, or n-butyl alcohol.

If, however, we start with a mixture of overall composition between B and C, the range of partial miscibility, such a mixture will be composed at the boiling point of two saturated solutions, B and C, will boil at a constant temperature, and will yield a vapor of constant composition, H, as long as the two layers are present. When the overall composition corresponds to H, obviously no separation is possible on distillation. When the overall composition is between C and H, however, the vapor is richer in water than the liquid phases, and hence the overall composition of the liquid phases must shift toward C on distillation. This behavior will result in the eventual disappearance of the layer containing butyl alcohol dissolved in water, leaving only a single layer of composition C. Further distillation of C will proceed along CD as described. On the other hand, when the overall starting composition is between B and H, removal of vapor richer in butyl alcohol than the liquid phases will result in a shift toward B, the alcohol-rich layer will disappear first, leaving only B, and further distillation of B will proceed along AB. It is seen, therefore, that overall compositions between A and H may be separated on complete distillation into a residue of pure A and a vapor of composition H, while overall compositions between H and D may be separated into a final distillate of H and a residue of pure D. No separation of H can be accomplished by distillation. In this respect the behavior of the system butyl alcohol-water is analogous to the behavior of minimum boiling point systems of completely miscible liquids. In fact, the latter type of diagram would result if points B and C were to approach each other and coincide with H.

The dotted lines emanating from B and C in Fig. 19(b) indicate the variation of the mutual solubility of the two liquids with temperature, and are part of the solubility diagram.

Other examples of systems exhibiting behavior similar to that of butyl

alcohol-water are aniline-water, *iso*-butyl alcohol-water, and ethyl acetate-water.

Vapor Pressure and Distillation of Immiscible Liquids. Since immiscible liquids are mutually insoluble, addition of one liquid to the other does not affect the properties of either liquid. Hence each will behave as if the other were not present. Consequently, in a mixture of two immiscible liquids each will exert the vapor pressure corresponding to the pure liquid at the given temperature, and the *total* vapor pressure above the mixture will be the sum of the vapor pressures of the two pure constituents, namely,

$$P = P_A^0 + P_B^0 \tag{11}$$

where P is the total vapor pressure, and P_A^0, P_B^0 are the vapor pressures of the two pure liquids A and B.

By definition the boiling point of any system is the temperature at which the total vapor pressure is equal to the confining pressure. Since the two liquids together can reach any given total pressure at a lower temperature than either liquid alone, it must follow that any mixture of two immiscible liquids must boil at a temperature *lower* than the boiling point of *either* of the two liquids. Furthermore, since at any given temperature there is no change in total vapor pressure with change in overall composition, the boiling point of all possible mixtures of the two must remain *constant* as long as the total pressure on the system does not change.

These facts are illustrated in Figs. 20(a) and 20(b), which give respectively the vapor pressure and boiling point diagrams for the immiscible system chlorbenzene-water. In Fig. 20(a) the horizontal line at P_B^0 gives the partial pressure of chlorbenzene
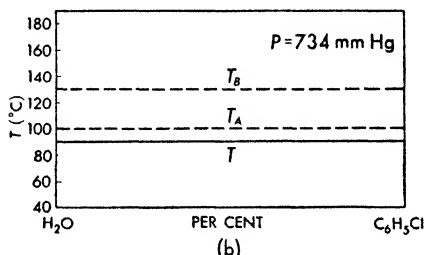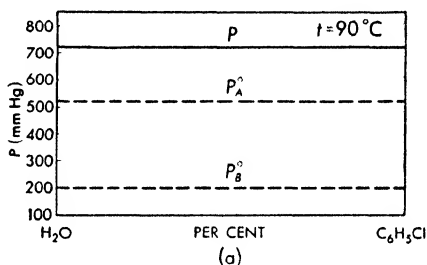
Fig. 20. Vapor Pressure and Distillation Diagrams for the System Water-Chlorbenzene

for all compositions at the given temperature, the line at P_A^0 the partial pressure of water, and the solid line the total vapor pressure above the two at the given temperature. Similarly in Fig. 20(b), the horizontal line at T_B gives the boiling point of chlorbenzene at the given total pressure, the line at T_A the boiling point of water, and the solid line the

boiling point of all possible mixtures of the two. The total vapor pressure is at all times constant, as is also the boiling point as long as both liquids are present. As soon as one of the liquids is boiled away, however, the boiling temperature will rise abruptly from that of the mixture to either T_A or T_B, depending on whether the chlorbenzene or the water is boiled away first.

At any boiling temperature for the mixture, T, the partial vapor pressures of the two constituents are P_A^0 and P_B^0 corresponding to the given temperature. If we let N_A and N_B be the mol fractions of the two constituents in the vapor, then $P_A^0 = N_A P$, $P_B^0 = N_B P$, and hence,

$$\frac{P_A^0}{P_B^0} = \frac{N_A P}{N_B P} = \frac{N_A}{N_B} \tag{12}$$

But $N_A = n_A/(n_A + n_B)$, and $N_B = n_B/(n_A + n_B)$, where n_A and n_B are the number of moles of A and B in any given volume of vapor. Consequently,

$$\frac{P_A^0}{P_B^0} = \frac{\dfrac{n_A}{n_A + n_B}}{\dfrac{n_B}{n_A + n_B}} = \frac{n_A}{n_B} \tag{13}$$

and, since the ratio of the partial pressures at T is constant, n_A/n_B must also be constant; i.e., the composition of the vapor is at all times constant as long as both liquids are present. Further, since $n_A = W_A/M_A$, and $n_B = W_B/M_B$, where W_A, W_B are the weights in any given volume and M_A, M_B are the molecular weights of A and B respectively, equation (13) becomes

$$\frac{P_A^0}{P_B^0} = \frac{n_A}{n_B} = \frac{\dfrac{W_A}{M_A}}{\dfrac{W_B}{M_B}}$$

$$= \frac{W_A}{W_B} \cdot \frac{M_B}{M_A}$$

$$\frac{W_A}{W_B} = \frac{M_A P_A^0}{M_B P_B^0} \tag{14}$$

Equation (14) relates directly the weights of the two constituents distilled from a mixture of two immiscible liquids to the molecular weights and the vapor pressures of the two pure constituents. It will be observed that the weight of any constituent distilled over depends on both its vapor pressure and molecular weight, and hence the effect of a low vapor

pressure tends to be counteracted by a high molecular weight as far as weight of a particular substance distilled over is concerned.

The facts, that a mixture of two immiscible liquids boils at a temperature lower than that of either alone, and that the weights of each obtained depend on both molecular weight and vapor pressure according to equation (14), are utilized industrially and in the laboratory for the purification of organic liquids which either boil at very high temperatures or tend to decompose when heated to their normal boiling point. The other liquid frequently is water, and the whole process is generally referred to as *steam distillation*. A mixture of the liquid and water, with which it is immiscible, is heated either directly or by injection of steam, and the vapors coming off are condensed and separated. In this manner it is possible to distill many liquids of high boiling point at temperatures below 100° C, the boiling point of water.

Distillation of immiscible liquids can also be utilized in determining the approximate molecular weight of one of the liquids involved. When the vapor pressures and weight ratios of distillates of two liquids are determined, and the molecular weight of one of the liquids is known, the molecular weight of the other can readily be calculated from equation (14). The manner in which the necessary data are obtained can best be explained through an example. When the two immiscible liquids chlorbenzene and water are boiled at a pressure of 734.4 mm Hg, the boiling point is 90° C, while the ratio of weight of chlorbenzene to water collected in the distillate is 2.47. Since at 90° C the vapor pressure of water is 526.0 mm Hg, the vapor pressure of chlorbenzene must be $734.4 - 526.0 = 208.4$ mm Hg. Therefore, letting A be chlorbenzene and B water, and applying equation (14),

$$M_A = \left(\frac{W_A}{W_B}\right)\left(\frac{P_B^0}{P_A^0}\right) M_B$$
$$= 2.47 \times \frac{526.0}{208.4} \times 18.02$$
$$= 112.3$$

The molecular weight of chlorbenzene calculated from atomic weights is 112.6.

REFERENCES FOR FURTHER READING

1. Badger and McCabe, "Elements of Chemical Engineering," second edition, McGraw-Hill Book Company, Inc., New York, 1936.
2. J. N. Friend, "A Textbook of Physical Chemistry," J. B. Lippincott Company, Philadelphia, 1933, Vol. I.
3. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.

4. J. H. Hildebrand, "Solubility of Non-Electrolytes," Reinhold Publishing Corporation, New York, 1936.
5. Hougen and Watson, "Industrial Chemical Calculations," second edition, John Wiley & Sons, Inc., New York, 1936.
6. Robinson and Gilliland, "Elements of Fractional Distillation," McGraw-Hill Book Company, Inc., New York, 1939. .
7. Walker, Lewis, and McAdams, "Principles of Chemical Engineering," third edition, McGraw-Hill Book Company, Inc., New York, 1937.

PROBLEMS

1. A solution contains 50% of water, 35% of ethyl alcohol, and 15% of acetic acid by weight. Calculate the mol fraction of each component in the mixture.
   Ans. $N_{H_2O} = 0.733$; $N_{HAc} = 0.066$; $N_{C_2H_5OH} = 0.201$

2. A Na_2CO_3 solution is made up by dissolving 22.5 g of $Na_2CO_3 \cdot 10\ H_2O$ in H_2O and adding H_2O until the total volume is 200 cc. The density of the resulting solution is 1.040 g/cubic centimeter. Calculate the molarity, normality, and mol fraction of Na_2CO_3 in the solution.
   Ans. 0.393 m; 0.786 N; 0.00704 mol fraction

3. A solution containing 10% of NaCl by weight has a density of 1.071 g/cubic centimeter. Calculate the molality and molarity of NaCl in the solution.

4. A gaseous solution was analyzed and found to contain 15% of H_2, 10% of CO, and 75% of N_2 by volume. What is the mol fraction and percentage by weight of each gas in the mixture?

5. Assuming dry air to contain 21% by volume of O_2, calculate the weight of O_2 dissolved in 1 liter of H_2O saturated with air under 1 atm pressure at 25° C. Use for Henry's law constant the average value computed from Table 1.
   Ans. 0.0085 g/liter

6. Using the solubility coefficient listed in Table 2, estimate the solubility of CO_2 in H_2O at 25° C under a partial pressure of CO_2 of 5 atm.

7. A mixture of $C_6H_5CH_3$ and C_6H_6 contains 30% by weight of $C_6H_5CH_3$. At 30° C the vapor pressure of pure $C_6H_5CH_3$ is 36.7 mm Hg while that of pure C_6H_6 is 118.2 mm. Assuming that C_6H_6 and $C_6H_5CH_3$ form ideal solutions, calculate the total pressure and the partial pressure of each constituent above the solution at 30° C.   *Ans.* $P_{C_6H_6} = 86.7$ mm; $P_{total} = 96.5$ mm

8. At 140° C the vapor pressure of C_6H_5Cl is 939.4 mm and that of C_6H_5Br is 495.8 mm. Assuming that these two liquids form an ideal solution, what will be the composition of a mixture of the two which boils at 140° C under 1 atm pressure? What will be the composition of the vapor at this temperature?

9. At 60° C the vapor pressure of pure ethyl alcohol is 352.7 mm Hg and that of methyl alcohol 625 mm Hg. A mixture of the two, which may be assumed to be ideal, contains 50% by weight of each constituent. What will be the composition of the vapor above the solution at 60° C?

10. At 30° C a mixture of C_6H_5OH and H_2O is made up 60% by weight of H_2O. The mixture splits into two layers, the C_6H_5OH layer containing 70% by weight of C_6H_5OH and the H_2O layer containing 92% by weight of H_2O. Calculate the relative weights of the two layers.   *Ans.* $W_{H_2O}/W_{Ph} = 0.934$

11. A mixture of aniline and H_2O is made up containing 30% by weight of aniline. From Fig. 16 determine graphically the relative weights of the two layers which form, and the proportion of the total amount of aniline present in each layer at 40 C.

12. Using data from a suitable handbook, plot on the same graph the vapor pressures of H_2O, C_6H_6, and the total pressure of a mixture of the two against temperature between 40 and 80° C. From this plot, find the boiling point of the immiscible system C_6H_6–H_2O under 1 atm pressure.

13. A totally immiscible liquid system composed of H_2O and an organic liquid boils at 90° C when the barometer reads 734 mm Hg. The distillate contains 73% by weight of the organic liquid. What is the molecular weight and vapor pressure at 90° C of the organic liquid? *Ans.* 122.9 g/mole; 208.2 mm Hg

14. Naphthalene may be steam distilled at 99.3° C under atmospheric pressure. What weight of steam will be required to carry 2 lb of naphthalene into the distillate at atmospheric pressure?

15. If the specific heat of steam is 0.5 cal/gram, and the heat of vaporization of the liquid in problem 13 is 78 calories/gram, calculate the total amount of heat and the minimum amount of steam, delivered at 99° C, required to steam-distill 500 g of the liquid at 90° C.

Colligative Properties of Solutions of Nonelectrolytes

In the preceding chapter were discussed the vapor pressure relations of binary solutions when both constituents of the solution are volatile. In this chapter we shall consider a number of properties of solutions which become manifest when one of the constituents of the binary solution is nonvolatile, or has such a low vapor pressure that it can be considered to be so. These properties are the vapor pressure lowering of the solvent, the freezing point lowering, the boiling point elevation, and the osmotic pressure of the solution. Taken together, these four properties of a solution are referred to as the *colligative properties*. A colligative property is any property which depends only on the *number* of particles in solution, and not in any way on the nature of these. As we shall see, this is the essential attribute of the four phenomena mentioned above, at least in dilute solutions.

The colligative properties of solutions deserve serious consideration, not only because they are of interest, but also because a study of these supplies valuable methods for estimating the molecular weight of the dissolved substance and evaluating a number of highly important thermodynamic quantities. In this chapter attention will be confined primarily to their utilization in evaluating the molecular weight of solutes in solutions of nonelectrolytes. In light of the principles to be developed in this chapter solutions of electrolytes behave "anomalously," and will be considered separately in the next chapter.

Lowering of Vapor Pressure of Solvent by Solutes. It has long been known that a nonvolatile solute lowers the vapor pressure of a liquid solvent in which it is dissolved. It is to Raoult (1887), however, that credit must be given for discovering the general principle underlying the effect of vapor pressure lowering and its dependence on the concentration of the solute in solution.

The vapor pressure lowering of the solvent due to solution in it of a nonvolatile solute can best be understood in terms of Raoult's law as developed in the preceding chapter. Consider a solution of a nonvolatile solute in a solvent. Let N_1 be the mol fraction of the solvent, N_2 the mol fraction of the solute, P^0 the vapor pressure of the pure solvent, and

P the vapor pressure of the solvent above the given solution. According to Raoult's law, then, P is given by

$$P = P^0 N_1 \qquad (1)$$

Since the solute is nonvolatile it does not contribute to the total vapor pressure, and hence equation (1) gives the total vapor pressure above the solution, which is due to solvent only. Furthermore, since N_1 is always less than unity, P must always be less than P^0. Consequently, *solution of a nonvolatile solute in a solvent will always result in lowering the vapor pressure of the solvent* below that of the pure solvent, and therefore the total vapor pressure above the solution will also always be lower than P^0. The extent of the vapor pressure lowering, ΔP, must be $P^0 - P$, or

$$\begin{aligned} \Delta P = P^0 - P &= P^0 - P^0 N_1 \\ &= P^0(1 - N_1) \\ &= P^0 N_2 \end{aligned} \qquad (2)$$

According to equation (2), the vapor pressure lowering of the solvent depends both on the vapor pressure of the solvent and the mol fraction of solute in solution. In other words, it depends on the nature of the solvent and on the concentration of solute, but not on the nature of the solute. However, if we consider the *relative vapor pressure lowering*, i.e., the ratio $\Delta P / P^0$, then from equation (2)

$$\frac{\Delta P}{P^0} = \frac{P^0 - P}{P^0} = N_2 \qquad (3)$$

and we see that the relative vapor pressure lowering of the solvent depends only on the mol fraction of solute, and is *completely independent* of either the nature of solute or solvent. Equation (3), the form Raoult's law takes for solutions of nonvolatile solutes, shows that the relative vapor pressure lowering of a solvent is a colligative property, because it depends only upon the concentration of the solute and upon nothing else.

The validity of Raoult's law applied to vapor pressure lowering may be tested either by comparing observed values of ΔP with those calculated from equation (2), or by comparing relative vapor pressure lowerings with the mol fraction of solute in solution according to equation (3). The first of these is used in Table 1, which shows observed and calculated vapor pressure lowerings for water at 20° C in solutions of mannite, while Table 2 gives a comparison of the relative vapor pressure lowerings of the solvent ether due to solution in it of various solutes of mol fraction N_2. Considering the difficulties involved in measuring small vapor pressure differences, the concordance between Raoult's law and experiment must be thought to be satisfactory.

TABLE 1

VAPOR PRESSURE LOWERING FOR AQUEOUS MANNITE SOLUTIONS AT 20° C
(P^0 = 17.51 mm Hg)

Moles of Mannite per 1000 g H_2O	ΔP Observed (mm Hg)	ΔP Calculated (mm Hg)
0.0984	0.0307	0.0311
0.1977	0.0614	0.0622
0.2962	0.0922	0.0931
0.4938	0.1536	0.1547
0.6934	0.2162	0.2164
0.8922	0.2792	0.2775
0.9908	0.3096	0.3076

TABLE 2

RELATIVE VAPOR PRESSURE LOWERING FOR VARIOUS SOLUTES IN ETHER SOLUTION

Solute	N_2	$\dfrac{P^0 - P}{P^0}$
Nitrobenzene	0.060	0.055
Ethyl benzoate	0.096	0.091
Methyl salicylate	0.092	0.086
Aniline	0.077	0.081
Benzaldehyde	0.130	0.132

Calculations Involving Raoult's Law. Equations (2) or (3) may be employed to calculate the vapor pressure lowering, and hence the vapor pressure, of solutions of nonvolatile solutes, or, knowing the vapor pressure lowering, they may be utilized to calculate the molecular weight of the dissolved substance. Consider the problem of calculating the vapor pressure of the solvent above a solution containing 53.94 grams of mannite (molecular weight = 182.11) per 1000 grams of water at 20° C. At this temperature the vapor pressure of water is 17.51 mm Hg. According to equation (2),

$$\Delta P = P^0 - P = P^0 N_2$$

$$= P^0 \left[\frac{\dfrac{W_2}{M_2}}{\dfrac{W_1}{M_1} + \dfrac{W_2}{M_2}} \right] \qquad (4)$$

where W_2 and M_2 are the weight and molecular weight respectively of the solute, while W_1 and M_1 are the same quantities for the solvent. Substituting the given data into equation (4), we obtain

$$\Delta P = 17.51 \left[\frac{\dfrac{53.94}{182.11}}{\dfrac{1000}{18.016} + \dfrac{53.94}{182.11}} \right]$$

$$= 17.51 \times 0.0053$$

$$= 0.0929 \text{ mm Hg}$$

The vapor pressure of the solution is, therefore,

$$P = P^0 - \Delta P = 17.51 - 0.09 = 17.42 \text{ mm Hg}$$

When the vapor pressure lowering for a particular solution of a solute in a solvent is known, the calculation may be reversed and equation (4) utilized to calculate M_2, the molecular weight of the solute — provided, of course, that the molecular weight of the solvent is known. For very dilute solutions the calculation can be simplified without serious error by omitting the term W_2/M_2 in the denominator. For such solutions W_2/M_2 is small compared to W_1/M_1, and equation (4) reduces to

$$\frac{P^0 - P}{P^0} = \frac{W_2 M_1}{W_1 M_2} \tag{5}$$

Determination of Vapor Pressure Lowering. The vapor pressure lowering of the solvent in solutions containing nonvolatile solutes may be ·measured by essentially the same methods as those described for the vapor pressure determination of pure liquids. These may be either static or dynamic in nature. The vapor need not be analyzed, as it must· be for solutions in which both constituents are volatile, for it is composed of solvent only. However, the solution composition in equilibrium with the vapor must be known and must be determined by some appropriate method.

These methods yield the vapor pressure of the solution directly. The vapor pressure lowering is then obtained by subtracting this pressure from that of the pure solvent at the same temperature. With concentrated solutions and with solvents of high volatility this procedure is fairly satisfactory. It leads to serious errors, however, when applied to dilute solutions in solvents of relatively low volatility, such as water at room temperature. This can be seen from the following example. The vapor pressure lowering to be expected for a $N_2 = 0.01$ solution of a solute in water at 20° C is $\Delta P = 17.51 \times 0.01 = 0.175$ mm Hg, and hence a vapor pressure for the solution of 17.33 mm. An uncertainty of $\pm$ 0.02 mm in the latter vapor pressure would constitute an error of only 0.11 per cent in the vapor pressure measurement, but this same uncertainty would affect ΔP by 11 per cent!

For such solutions it is preferable to employ one of a number of *differential* methods which·yield directly the vapor pressure lowering of the solvent in any particular solution. The simplest of these involves the use of the Frowein tensimeter. This tensimeter consists essentially of a U-tube manometer with two flasks attached to its upper ends. In one of these flasks is placed the pure solvent, while the other holds the solution whose vapor pressure lowering is to be determined. As a mano-, metric fluid, some liquid of very low volatility and relatively low density

is used, for example, β-bromonaphthalene, n-dibutyl phthalate, or Apiezon oil. To remove any dissolved air the two vessels are first connected together, and the whole apparatus is thoroughly evacuated. The two vessels are then isolated from each other, the apparatus is placed in a constant temperature bath, and, after establishment of equilibrium, the difference in pressure between the two sides is read on the manometer. This difference is directly the vapor pressure lowering of the solvent in the given solution at the temperature of the thermostat. With careful attention and precautions an accuracy of 0.01 mm Hg can be attained.

Boiling Point Elevation of Solutions. Solutions containing nonvolatile solutes boil at temperatures *higher* than the boiling point of the pure solvent. The difference between the boiling points of the solution and pure solvent at any given constant pressure is referred to as the *boiling point elevation* of the solution. Observation has shown that the boiling point elevation of a solution depends on the nature of the solvent and the concentration of solute, but is independent, at least in dilute solutions, of the nature of the solute as long as the latter is not ionized.
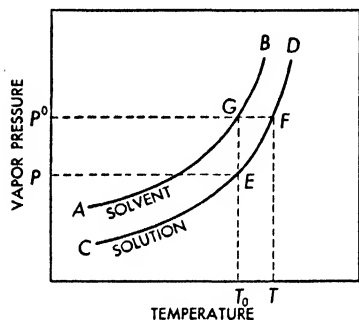
Fig. 1. Boiling Point Elevation Due to Solutes

This elevation of the boiling point is readily understood in terms of the vapor pressure lowering, and is a direct consequence of it. Consider the vapor pressure-temperature diagram shown in Fig. 1. In this diagram curve AB represents the vapor pressure of the pure solvent as a function of temperature. Since the vapor pressure of the solution is at all temperatures lower than that of the solvent, the vapor pressure-temperature curve of the solution must lie below that of the pure solvent, and hence must be represented by some curve such as CD in the figure. In order to reach the boiling point at some given external pressure P^0, the solvent and the solution must be heated to temperatures at which their respective vapor pressures become equal to the given confining pressure. As the diagram indicates, the solvent can attain the pressure P^0 at the temperature T_0, but the solution must be raised to temperature T, higher than T_0, before the same pressure is reached. Consequently, at the same external pressure the solution must boil at a temperature higher than the pure solvent; and, the elevation of the boiling point of the solution, ΔT_b, is given by $\Delta T_b = T - T_0$. These considerations are perfectly general, and apply to any solution of a nonvolatile solute.

Dependence of ΔT_b on Concentration. By applying the Clausius-Clapeyron equation and Raoult's law to the conditions depicted in Fig. 1, it is readily possible to deduce a relation between the boiling point elevation of the solution and its concentration. Since points E and F both lie on the vapor pressure curve of the solution, they are both given by the Clausius-Clapeyron equation,

$$\ln \frac{P^0}{P} = \frac{\Delta H_v}{R}\left(\frac{T - T_0}{TT_0}\right) \tag{6}$$

where P is the vapor pressure of the solution at temperature T_0, while P^0 is the vapor pressure at temperature T. ΔH_v is heat of vaporization of the solvent per mole from the solution. When the solution is dilute, this is essentially the heat of vaporization per mole of the pure solvent. Again, when the solution is dilute T is not much different from T_0, and hence we may write $TT_0 = T_0^2$. Consequently equation (6) becomes

$$\ln \frac{P^0}{P} = -\ln \frac{P}{P^0} = \frac{\Delta H_v}{R} \cdot \frac{\Delta T_b}{T_0^2} \tag{7}$$

In equation (7) P is the vapor pressure of the solution at T_0, while P^0 is also the vapor pressure of the pure solvent at the same temperature. When Raoult's law is applicable to the solution, these two pressures are related through

$$\frac{P}{P^0} = N_1 = 1 - N_2 \tag{8}$$

where N_2 is the mol fraction of solute in the solution, and hence equation (7) becomes

$$\ln\,(1 - N_2) = -\frac{\Delta H_v}{R} \cdot \frac{\Delta T_b}{T_0^2} \tag{9}$$

An expansion of $\ln\,(1 - N_2)$ in series yields the expression,

$$\ln\,(1 - N_2) = -N_2 - \frac{N_2^2}{2} - \frac{N_2^3}{3} + \cdots$$

and since the solution was already specified to be dilute, N_2 must be small, and all terms beyond the first in the expansion can be considered negligible. Writing, then, $-N_2$ for $\ln\,(1 - N_2)$ in equation (9), we have

$$-N_2 = -\frac{\Delta H_v}{R} \cdot \frac{\Delta T_b}{T_0^2}$$

and therefore,

$$\Delta T_b = \frac{RT_0^2}{\Delta H_v} \cdot N_2 \tag{10}$$

Equation (10) gives the boiling point elevation of a solution in terms of the boiling point and heat of vaporization of the solvent, and the mol fraction of the solute in solution. Since for any given solvent T_0 and ΔH_v are constant, the boiling point elevation for dilute solutions is seen to be directly proportional to the mol fraction of the solute only, and is in no way dependent on the nature of the solute. The boiling point elevation of a solution is thus a colligative property.

The common practice in boiling point elevation work is to express the concentration not in mol fractions but in *moles of solute per 1000 grams of solvent*. If we let m be the number of moles of solute dissolved in 1000 grams of solvent, and n_1 be the number of moles of solvent in 1000 grams, then,

$$N_2 = \frac{m}{n_1 + m} = \frac{m}{n_1}$$

since for dilute solutions m is small compared to n_1 and may be disregarded. Therefore,

$$\Delta T_b = \left[\frac{RT_0^2}{\Delta H_v n_1} \right] m \tag{11}$$

For any given solvent all the quantities in the brackets of equation (11) are constant, and hence the whole term is constant. Writing

$$K_b = \frac{RT_0^2}{\Delta H_v n_1} \tag{12}$$

equation (11) finally reduces to

$$\Delta T_b = K_b m \tag{13}$$

According to equation (13) the boiling point elevation of any dilute solution is directly proportional to the number of moles of solute per 1000 grams of solvent; i.e., *the molality of the solution.* The proportionality constant K_b is called either the *molal boiling point elevation constant,* or *the ebullioscopic constant,* and signifies the rise in boiling point for a 1 molal solution of a solute in a solvent provided the laws of dilute solutions are applicable to such a concentration. Actually it is the boiling point elevation per mole calculated by proportion from the boiling point elevation of much more dilute solutions.

Validity of Equation (13). The validity of equation (13) deduced theoretically for dilute solutions on the assumption of Raoult's law may be tested in a number of ways. First, the equation demands that for a given solvent the boiling point elevation be proportional to the molality of the solute irrespective of its nature. Second, the equation demands

that for any given solvent the proportionality constant K_b be independent of the nature or concentration of the solute. In both these respects the equation is in good agreement with experimental results on dilute solutions as long as the solutes are nonelectrolytes. A third, and more crucial, test can be made by comparing the observed values of K_b with those predicted by equation (12). If equation (12) is valid, it should be possible to calculate the molal boiling point elevation constant from a knowledge of the normal boiling point and heat of vaporization of the solvent. For instance, for water $T_0 = 373.2°$ K, while the heat of vaporization at the boiling point is 539 calories per gram. Applying equation (12), the molal boiling point elevation constant for water as a solvent follows as

$$K_b = \frac{RT_0^2}{\Delta H_v n_1}$$
$$= \frac{1.987 \times (373.2)^2}{(18.02 \times 539)\dfrac{(1000)}{18.02}}$$
$$= 0.513°$$

This value compares very well with the experimentally observed $K_b = 0.52°$.

Table 3 lists the boiling points and observed ebullioscopic constants for a number of solvents. The molal elevation constants for these solvents calculated by equation (12) are also included for comparison. The agreement between calculated and observed values of this constant is highly satisfactory when uncertainties involved in boiling point determinations are considered. In view of this concordance between theory and experiment, equation (12) may be employed also to evaluate heats of vaporization of the solvent from experimental values of K_b.

TABLE 3

MOLAL BOILING POINT ELEVATION CONSTANTS

Solvent	Boiling Point (° C)	K_b (obs.)	K_b (calc.)
Acetone	56.5	1.72	1.73
Carbon tetrachloride	76.8	5.0	5.02
Benzene	80.1	2.57	2.61
Chloroform	61.2	3.88	3.85
Ethyl alcohol	78.4	1.20	1.19
Ethyl ether	34.6	2.11	2.16
Methyl alcohol	64.7	0.80	0.83
Water	100.0	0.52	0.51

Determination of Boiling Point Elevation. If we let ΔT_b be the boiling point elevation for a solution containing w_2 grams of solute of

molecular weight M_2 dissolved in w_1 grams of solvent, then the weight of solute per 1000 grams of solvent is

$$\frac{w_2 \times 1000}{w_1}$$

and hence m, the molality of the solution, is

$$m = \frac{w_2 \times 1000}{w_1 M_2} \tag{14}$$

Equation (13) in terms of equation (14) is, therefore,

$$\Delta T_b = K_b \left(\frac{1000\, w_2}{w_1 M_2} \right) \tag{15}$$

In order to obtain K_b from this relation it is necessary to measure ΔT_b, w_1, and w_2 with a solute of known molecular weight. On the other hand, in order to determine M_2, it is necessary to ascertain ΔT_b, w_1, and w_2 for a solvent of known K_b. In both instances the data to be obtained are the weight of solute, w_2, the weight of solvent, w_1, and the boiling point elevation of the solution, ΔT_b. A method for obtaining this information will now be described.

In Fig. 2 is shown the Cottrell apparatus for determining boiling point elevations. It consists of a vessel A fitted with a vapor condenser D and a ground glass joint from which extends a glass shield B, whose function is to prevent the condensed vapors from D from coming into contact with the thermometer. The thermometer itself is suspended in the vapor above the liquid and inside the shield. To assure equilibrium between vapor and boiling liquid at the thermometer bulb, a "siphon pump" C is incorporated into the apparatus. This "pump" consists of a glass tube having a funnel-shaped base and a bent U-shaped top which fits around the thermometer stem as shown. When the liquid is boiling, bubbles of vapor formed at the bottom rise upward and force any

Fig. 2. Cottrell Boiling Point Apparatus

liquid trapped in the tube to be spilled over the thermometer bulb. By this simple device the thermometer is always bathed by a mixture of boiling liquid and vapor, and reaches a steady equilibrium temperature in a very short time. The liquid may be boiled either with a small gas flame or electrically.

For a determination, a weighed quantity of solvent, w_1, is introduced into the boiling tube. Heating is then started and continued until the solvent boils uniformly and the temperature becomes constant. This temperature is read and is the boiling point of the solvent. Heating is now stopped, a weighed quantity, w_2, of solute pressed into a pellet is introduced, and the heating is repeated until a constant boiling point is again attained. The reading on the thermometer at this point minus the reading for the boiling point of the pure solvent gives the boiling point elevation ΔT_b of the solution.

Among a number of other types of apparatus described in the literature may be mentioned those of Beckmann, McCoy, Davis and Brandt, Saxton and Smith, and Swietoslawski. The various modifications in design and technique have all been aimed at the prevention of superheating of the liquid, the more rapid attainment of equilibrium, and greater accuracy of temperature and concentration determination. Some of the changes introduced have involved the use of manostats for pressure control on the boiling system, the use of thermocouples for temperature measurement, and the use of dual vessels for the simultaneous ebullition of solvent and solution.

Calculation of Molecular Weights from Boiling Point Elevation. When the ebullioscopic constant of a solvent is known, a determination of the boiling point elevation of a solution containing the unknown solute in definite concentration is sufficient to yield the molecular weight of the solute. When the ebullioscopic constant is unknown, however, an independent determination of ΔT_b must be made with a solute of known molecular weight. The calculations involved can best be seen from the following example. A solution containing 0.5126 gram of naphthalene (molecular weight = 128.16) in 50.00 grams of carbon tetrachloride yields a boiling point elevation of 0.402° C, while a solution of 0.6216 gram of an unknown solute in the same weight of solvent gives a boiling point elevation of 0.647° C. Find the molecular weight of the unknown solute.

For finding K_b of carbon tetrachloride, the given data are

$$w_1 = 50.00 \text{ g} \qquad \Delta T_b = 0.402° \text{ C}$$
$$w_2 = 0.5126 \text{ g} \qquad M_2 = 128.16$$

Substituting these values into equation (15) and solving for K_b, we obtain

$$K_b = \frac{\Delta T_b \, w_1 M_2}{1000 \, w_2}$$
$$= \frac{0.402 \times 50.00 \times 128.16}{1000 \times 0.5126}$$
$$= 5.03° \text{ C/mole/1000 g of solvent}$$

Using now the found value of K_b along with the information for the unknown,

$$w_1 = 50.00 \text{ g} \qquad \Delta T_b = 0.647$$
$$w_2 = 0.6216 \text{ g}$$

equation (15) yields, for M_2 of the unknown,

$$M_2 = \frac{1000 \, w_2 K_b}{w_1 \cdot \Delta T_b}$$
$$= \frac{1000 \times 0.6216 \times 5.03}{50.00 \times 0.647}$$
$$= 96.7$$

Freezing Point Lowering of Solutions. When a dilute solution of a solute in a solvent is cooled, a temperature is eventually reached at which *solid solvent* begins to separate from solution. The temperature at which this separation begins is called *the freezing point of the solution*. More generally the freezing point of a solution may be defined as the temperature at which a particular solution is in equilibrium with solid solvent.

Solutions freeze at *lower* temperatures than the pure solvent. The fact that solution of a solute in a solvent lowers the freezing point of the solvent has been known for a long time, and the dependence of this lowering on concentration was studied as early as the latter half of the eighteenth century by Watson and Blagden. However, Beckmann, starting in 1888, is the man who placed the subject of *cryoscopy*, the study of freezing points of solutions, on a solid experimental foundation, while the theoretical treatment of freezing point depression is due primarily to J. H. van't Hoff (1886).

The freezing point lowering of a solution, like the boiling point elevation, is a direct consequence of the vapor pressure lowering of the solvent by dissolved solute. To appreciate this, consider the vapor pressure-temperature diagram shown in Fig. 3. In this diagram AB is the sublimation curve of the solid solvent, while CD is the vapor pressure curve of pure liquid solvent. At the freezing point of the pure solvent the solid and liquid phases are in equilibrium, and consequently they must both have at this temperature identical vapor pressures. The only point on the diagram at which the two forms of the pure solvent have the same

vapor pressure is B, the intersection of AB and CD, and therefore T_0, the temperature corresponding to B, must be the freezing point of the pure solvent. When a solute is dissolved in the solvent, however, the vapor pressure of the latter is lowered below that of the solid, and equilibrium can no longer exist at T_0. To ascertain the new point of equilibrium between the solution of the solute and the solid solvent, the
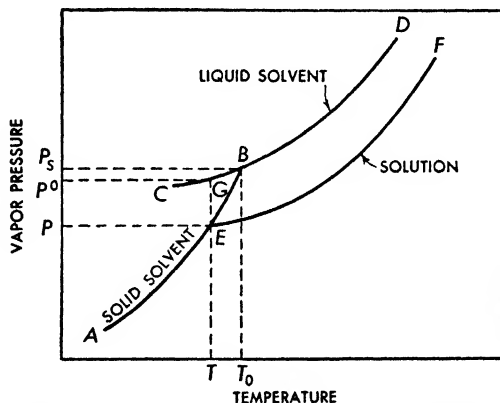
Fig. 3. Depression of Freezing Point by Solutes

temperature must be found at which the vapor pressure of the solution becomes equal to that of the solid, i.e., at which the vapor pressure curve of the solution intersects the sublimation curve, and this will be the freezing point of the solution. Since the vapor pressure curve of the solution, EF, always lies below that of the pure solvent, the intersection of EF and AB can occur only at a temperature lower than T_0. Hence any solution of the solute in the solvent must have a freezing point, T, lower than that of the solvent, T_0.

Dependence of Freezing Point Depression on Concentration. The freezing point depression of a solution, ΔT_f, is defined as $\Delta T_f = T_0 - T$, and represents the number of degrees by which the freezing point of a solution is lower than that of the pure solvent. Experiment has established that the magnitude of ΔT_f for a solution depends both on the nature of the solvent and the concentration of the solution. For dilute solutions of various solutes in a given solvent ΔT_f varies linearly with concentration irrespective of the nature of the solute — as long as it is not ionized. The proportionality constant of this concentration variation is, however, a function of the solvent, and varies considerably for different solvents.

To relate mathematically the freezing point depression of a solution to the factors mentioned, consider again Fig. 3. Let P_s be the vapor pressure of solid and pure liquid solvent at T_0, and P be the vapor pressure

of both solid solvent and solution at temperature T. Again, let P^0 be the vapor pressure of pure supercooled liquid solvent at T, point G. Then, since points G and B lie on the same vapor pressure curve, they must both be related by the Clausius-Clapeyron equation,

$$\ln \frac{P_s}{P^0} = \frac{\Delta H_v(T_0 - T)}{RT_0T} \tag{16}$$

where ΔH_v is the heat of vaporization of the pure solvent. Similarly, since points E and B lie on the same sublimation curve, they must be given by the equation

$$\ln \frac{P_s}{P} = \frac{\Delta H_s(T_0 - T)}{RT_0T} \tag{17}$$

where ΔH_s is the heat of sublimation of the solid solvent. Subtracting now equation (17) from equation (16), we obtain

$$\ln P_s - \ln P^0 - \ln P_s + \ln P = \frac{\Delta H_v(T_0 - T)}{RT_0T} - \frac{\Delta H_s(T_0 - T)}{RT_0T}$$

$$\ln P - \ln P^0 = -\frac{(\Delta H_s - \Delta H_v)(T_0 - T)}{RT_0T}$$

$$\ln \frac{P}{P^0} = -\frac{(\Delta H_s - \Delta H_v)(T_0 - T)}{RT_0T} \tag{18}$$

But $(\Delta H_s - \Delta H_v) = \Delta H_f$, the heat of fusion of the solvent. Therefore,

$$\ln \frac{P}{P^0} = -\frac{\Delta H_f(T_0 - T)}{RT_0T} = -\frac{\Delta H_f \Delta T_f}{RT_0T} \tag{19}$$

Equation (19) relates the vapor pressure of solid solvent at temperature T to the vapor pressure of pure liquid solvent at the same temperature. But since the vapor pressures of solid solvent and solution are equal at temperature T, the freezing point of the solution, equation (19) also relates the vapor pressure of the *solution* to that of the pure solvent at the temperature T. If we assume now that Raoult's law is applicable to the solution, then $P/P^0 = N_1 = (1 - N_2)$, where N_1 and N_2 are the mol fractions of the solvent and solute in solution, and equation (19) becomes

$$\ln (1 - N_2) = -\frac{\Delta H_f \Delta T_f}{RT_0T} \tag{20}$$

When N_2 is small, i.e., when the solution is dilute, $\ln (1 - N_2)$ is equal essentially to $- N_2$ and T_0T to T_0^2. Hence,

$$- N_2 = - \frac{\Delta H_f \Delta T_f}{R T_0^2}$$

$$\Delta T_f = \left(\frac{R T_0^2}{\Delta H_f} \right) N_2 \tag{21}$$

Finally, designating by m the *molality* of the solution, by n_1 the number of moles of solvent in 1000 grams, $N_2 = m/n_1$ (approximately) and

$$\Delta T_f = \left(\frac{R T_0^2}{\Delta H_f \, n_1} \right) m$$

$$= K_f \, m \tag{22}$$

where

$$K_f = \frac{R T_0^2}{\Delta H_f \, n_1} \tag{23}$$

Equation (22) is the fundamental relation of cryoscopy, and is directly analogous to equation (13) for boiling point elevation. K_f, called the *molal freezing point lowering* or *cryoscopic constant* of a solvent, is defined in terms of quantities characteristic of the solvent only, and in no way depends on either the concentration or nature of the solute. It is the freezing point depression counterpart of the boiling point elevation constant K_b. Since for any given solvent K_f is a constant, the freezing point depression of a solution is determined by the concentration of solute only, and hence the freezing point depression, like the vapor pressure lowering and boiling point elevation, is a colligative property.

Validity of Equation (22). An important test of this equation is the constancy of K_f with concentration for a given solvent. Table 4 shows some experimental data for the freezing point lowering of solutions of urea in water, and the value of $K_f = \Delta T_f / m$ calculated therefrom. As the last column indicates, K_f is constant throughout the given concentration range and equal to 1.85. Essentially identical values for the cryoscopic constant of water have been obtained with numerous other solutes in water. Furthermore, the observed value of this constant is in good agreement with the value predicted for water as a solvent by equation (23). For water, $T_0 = 273.2$, $\Delta H_f = 79.71 \times 18.02$, and $n_1 = 1000/18.02$. Therefore,

$$K_f = \frac{1.987(273.2)^2}{(79.71 \times 18.02) \left(\dfrac{1000}{18.02} \right)}$$

$$= 1.857° \text{ C/mole/1000 grams } H_2O$$

Similar concordance between theory and experiment has been obtained with many other solvents. Table 5 lists a number of solvents, their freezing points, and their cryoscopic constants as either evaluated experi-

mentally or calculated from (23). It will be observed that, of those listed, the cryoscopic constant of water is the lowest. Hence, for a given concentration of solute, more pronounced freezing point depressions may be obtained in the various other solvents, and this fact is of importance in the practical use of these solvents for molecular weight determination.

It should be emphasized that because of the assumptions made in its derivation, equation (22) can be expected to be valid only in dilute solutions. In more concentrated solutions K_f deviates considerably from constancy, and the deviation is the more pronounced the more concentrated the solution. Thus the cryoscopic constant of benzene as evaluated from freezing point determinations of carbon tetrachloride in the solvent is $5.09°$ C per mole in a 0.1184 molal solution, and only $4.82°$ C per mole in a 1.166 molal solution.

<div align="center">

TABLE 4

FREEZING POINT DEPRESSIONS FOR SOLUTIONS OF UREA IN WATER

</div>

m	ΔT_f	$K_f = \dfrac{\Delta T_f}{m}$
0.000538	0.001002	1.862
0.004235	0.007846	1.851
0.007645	0.01413	1.849
0.012918	0.02393	1.850
0.01887	0.03496	1.853
0.03084	0.05696	1.848
0.04248	0.07850	1.848
		1.852

<div align="center">

TABLE 5

CRYOSCOPIC CONSTANTS FOR VARIOUS SOLVENTS

</div>

Solvent	Freezing Point (°C)	K_f
Acetic acid	16.7	3.9
Benzene	5.5	5.12
Bromoform	7.8	14.4
Camphor	178.4	37.7
Cyclohexane	6.5	20.0
1,4-Dioxane	10.5	4.9
Naphthalene	80.2	6.9
Phenol	42	7.27
Tribromophenol	96	20.4
Triphenylphosphate	49.9	11.76
Water	0.00	1.86

Determination of Freezing Point Lowering. Of the various methods developed for freezing point lowering measurements, only three will be described, namely: (1) the Beckmann method, (2) the Rast method, and (3) the equilibrium method.

(1) The Beckmann Method. A simple Beckmann freezing point apparatus is shown in Fig. 4. The freezing point tube *A* has a side arm *B*, for introduction of solute, and is fitted with a stopper carrying a Beckmann thermometer and stirrer. To prevent too rapid cooling of the contents of the freezing tube, *A* is surrounded with a guard tube *C*, as indicated, so as to leave an air space between *A* and *C*. The whole assembly is supported in the large beaker *D*, which contains the cooling mixture. This mixture should be at a temperature no lower than 5° below the freezing point of the solvent used.

For a determination, a definite weight of solvent is placed in *A* and cooled with stirring to a temperature about 0.5° below the freezing point. Rapid stirring of the supercooled liquid initiates the crystallization, and the temperature rises rapidly to the freezing point, which is recorded. Tube *A* is now removed, warmed to melt the solid, and a definite weight of solute added through *B*. After the solute has dissolved, *A* is replaced, and the freezing point is redetermined in exactly the same manner as for the solvent. The difference in the two freezing points is the freezing point lowering of the solution.

(2) Rast Method. The Rast method is particularly convenient for molecular weight determinations of solutes that are soluble in

Fig. 4. Beckmann Freezing Point Apparatus

camphor. It takes advantage of the fact that the cryoscopic constant of camphor is very high, 37.7° C per mole per 1000 grams of solvent. The Rast procedure is essentially a micromethod, for a freezing point lowering determination can be made with only a few milligrams of solute. Some solid camphor is first dusted into a small capillary tube, and a melting point determination is made by the same technique as is ordinarily employed in organic chemistry. A solution of the solute in camphor is prepared, then, as follows: A small weight of solute and about 10 to 15 times that weight of camphor are mixed and melted by heating over a flame. The melt is permitted to solidify, the solid is ground to a fine powder, and a small quantity of the powder is transferred into a clean capillary tube. The tube is heated now slowly until the solid begins to melt, and the temperature at which the last crystals of camphor disappear is taken as the freezing point of the solution. Since the lowerings obtained with
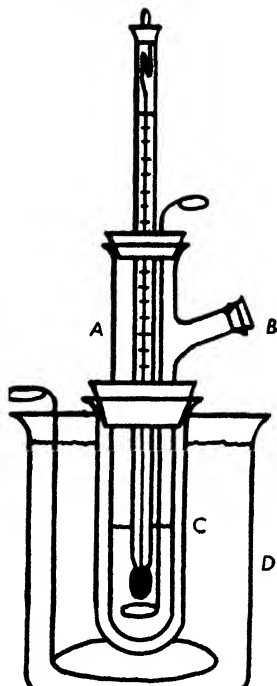
this method are usually of the order of 10° or more, an ordinary thermometer whose readings can be estimated to 0.1 or 0.2° is satisfactory for the accuracy required in molecular weight determinations.

(*3*) *Equilibrium Method.* The Rast method of freezing point lowering determination is limited to camphor or any other solvent having a high cryoscopic constant, such as pinene dichloride ($K_f = 56.2$) or pinene dibromide ($K_f = 80.9$). It is suitable only for the approximate estimation of molecular weights of solutes soluble in such solvents. Although the Beckmann method is of more general applicability, it is open to a number of objections. First, since most solvents and solutions exhibit a ready tendency to supercool, there is always the danger that the observed freezing points are not the true equilibrium temperatures. Second, the crystallization of any appreciable amount of solvent changes the concentration of solution, and hence the concentrations of solution calculated from the weights are always too low. For these reasons the Beckmann method is rarely used now for the precise determination of freezing point depression.

Practically all accurate freezing point lowering data are obtained at present by the equilibrium method. In this method an intimate mixture of solid and liquid solvent is first prepared, and then agitated until equilibrium between the two is established. The temperature at this point is read, usually with a multijunction thermocouple. A volume of concentrated solution of the solute in the solvent is now added to yield approximately some desired concentration, and equilibrium is again permitted to be reestablished. The temperature is then read, and a sample of the liquid phase is removed and analyzed. By this procedure the tendency to supercool is reduced to a minimum, and the concentration of solution obtained is that at equilibrium.

The method can be improved even further to yield directly the freezing point depression of the solution. How this is accomplished can be gathered from the apparatus (Fig. 5) used by L. H. Adams [1] for the measurement of freezing point lowering of aqueous solutions. Two similar Dewar vessels, *A* and *B*, were enclosed in a felt-covered box, and then packed on all sides with crushed ice. About 400 grams of pure ice and 600 grams of pure water were placed in vessel *A*, and the same weight of ice and solution in *B*. The contents of the two vessels were agitated with the two special stirrers *S* and *S'* which operated alternately with an up and down motion. These stirrers were essentially lift pumps which on the up stroke lifted the liquid and spilled it over the ice. The difference in temperature between the two vessels was measured by means of a multijunction thermoelement *T*, one end of which was immersed in *A*, the other in *B*. When equilibrium was established, as

[1] L. H. Adams, J. Am. Chem. Soc., **37**, 481 (1915).

determined by the constancy of the temperature difference between the two vessels, a sample was removed from B with the pipette P, and the solution was analyzed for solute content. With this scheme Adams was
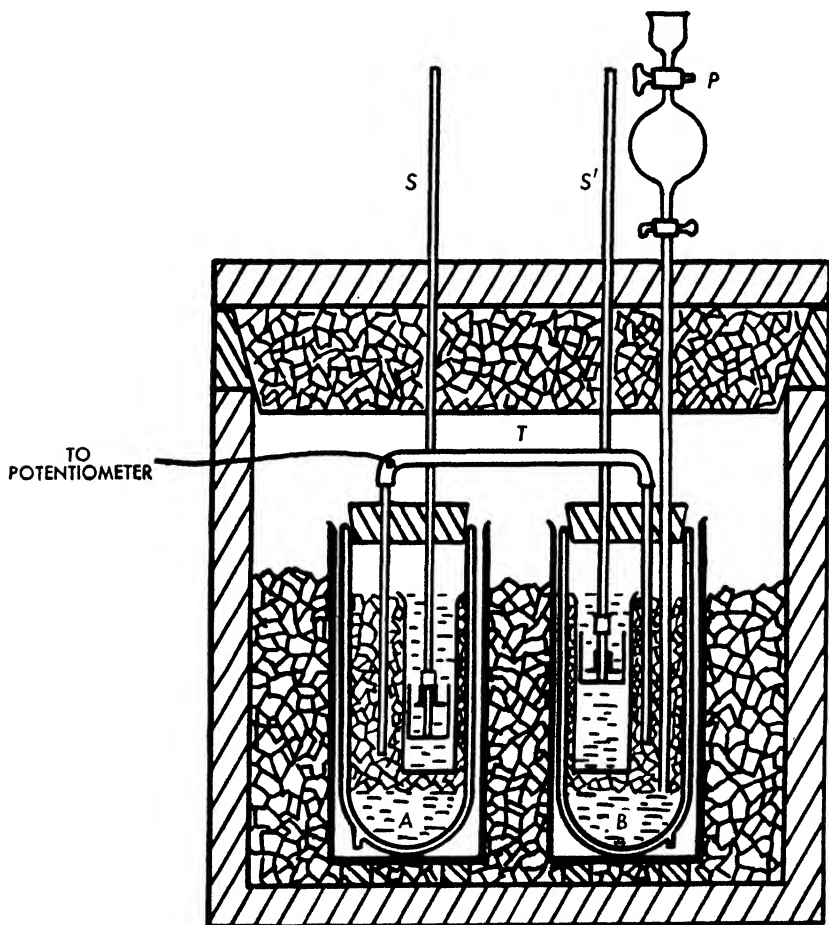
Fig. 5. Adams's Apparatus for Freezing Point Depression

able to determine temperature differences of $0.0001°$ C. In some other researches of similar nature precisions of 0.00001 to $0.00002°$ C have been reported.

The accuracies that have been attained in freezing point lowering work are considerably higher than any that have been possible in the measurement of the other colligative properties. For this reason freezing point depressions of solutions have been determined more extensively than either vapor pressure lowering, boiling point elevation, or, as we shall see, osmotic pressure, and are used more frequently than the others

in evaluating certain important thermodynamic quantities. Because of the difficult techniques involved, osmotic pressure determinations are not made very often, and they usually are the least accurately determined of the colligative properties. One point which militates against both the boiling point elevation and freezing lowering is the fact that any data obtained apply only to the temperatures at either the boiling point or freezing point of the solution. In this respect vapor pressure lowerings and osmotic pressures are more flexible, for they can be measured at any temperature between the freezing point and boiling point of the solvent. However, experimental methods developed recently permit the measurement of boiling point elevation of solutions at various reduced pressures, and thus make possible the extension of the temperature range over which this property can be estimated.

Calculation of Molecular Weights from Freezing Point Lowering. In view of the ease with which fairly precise freezing point data can be obtained, such data are particularly suitable for determining molecular weights of solutes in solution. The calculations involved are exactly analogous to those made in conjunction with boiling point elevation, as are also the equations and data required. If, again, we designate by w_2 the weight of solute of molecular weight M_2 dissolved in w_1 grams of solvent, then, according to equation (14),

$$m = \frac{w_2 \times 1000}{w_1 M_2}$$

and hence equation (22) becomes

$$\Delta T_f = K_f\, m$$
$$= K_f \frac{(1000\, w_2)}{w_1 M_2} \tag{24}$$

From this equation the molecular weight follows as

$$M_2 = K_f \left(\frac{1000\, w_2}{\Delta T_f\, w_1} \right) \tag{25}$$

Therefore, to calculate the molecular weight, K_f for the solvent must be known, and ΔT_f, w_1, and w_2 must be measured. If K_f is not known, it can be either calculated by means of equation (23), or determined by making first a freezing point determination with a solution of a solute of known molecular weight in the same solvent.

Separation of Solid Solutions on Freezing. These freezing point depression considerations and the equations deduced are valid only when the solid separating from solution is the pure solvent. Occasionally cases are encountered, such as solutions of iodine in benzene or thiophene in

benzene, where the solid crystallizing out contains solute dissolved in it in the form of solid solutions. For such cases equation (22), and all others based on it, are no longer applicable. It can be shown by theoretical argument that when the solute crystallizing from solution is a solid solution, equation (22) must be replaced by

$$\Delta T_f = K_f(1 - k)m \tag{26}$$

where the symbols have the same significance as before, and k is the ratio of mol fraction of solute in the solid to mol fraction of solute in solution. When the solid phase is pure solvent, k is zero, and equation (26) reduces to equation (22). When the solid phase is not pure, however, two conditions may be distinguished. If the solute is more soluble in the liquid solvent than in the solid, k is a positive fraction, and hence $(1 - k)$ is less than unity. The effect will be, therefore, to give a freezing point depression less than would be anticipated for separation of pure solid solvent. If, on the other hand, the solute is more soluble in the solid phase than in the liquid, k is a positive number greater than unity, and $(1 - k)$ is negative. Under these conditions ΔT_f is also *negative;* i.e., a freezing point *elevation* rather than a lowering is observed for the solution. Such behavior is rarely encountered under ordinary circumstances, but it is not at all unusual with metals and salts that form solid solutions.

Osmosis and Osmotic Pressure. When an aqueous solution of a solute is separated from pure water by a semipermeable membrane, i.e., a membrane that permits the passage of water but not of dissolved solute, the water always tends to pass through the membrane into the solution, diluting it. The phenomenon, called *osmosis*, was first reported by the Abbé Nollet (1748) and was further investigated by Dutrochet (1827–1832) and Vierordt (1848). These men all employed animal membranes. Graham (1854) showed, however, that such membranes are not truly semipermeable, and this fact led Traube (1864) to investigate the possibility of developing an artificial membrane which would more closely approximate the condition of semipermeability. He found that the membrane most satisfactory in this respect was a thin film of copper ferrocyanide, $Cu_2Fe(CN)_6$, prepared by contacting a solution of a cupric salt with one of potassium ferrocyanide. But it is to the botanist Pfeffer (1877) that credit must be given for improving the strength and quality of copper ferrocyanide membranes, and for being the first to make quantitative measurements of the osmotic pressure of solutions.

Before proceeding to a discussion of Pfeffer's work or any of the subsequent refinements, it is necessary to define the measure of the tendency to osmose, namely, the *osmotic pressure*. For this purpose consider the diagram shown in Fig. 6. In this diagram is a chamber A, open at one

end and fitted at the other with a movable piston B. Further, the chamber is divided by means of a semipermeable membrane C into two sections, of which the right one is filled with pure water, the other with some aqueous solution. Because of osmosis, water will tend to pass through the membrane into the solution and displace the piston upward. The motion of the piston upward, and hence osmosis of water, can be prevented, however, by the application of pressure to the piston in such a way that it remains stationary in its original position. The mechanical pressure which must be applied on a solution to prevent osmosis of the solvent into the solution through a semipermeable membrane is called the *osmotic pressure* of the solution. This pressure for a given solution depends on a number of factors, as we shall see later, but it does not depend on the nature of the membrane so long as the membrane is truly semipermeable. Hence the osmotic pressure of a solution must be considered as a measure of some real difference, expressible in pressure units, in the natures of the pure solvent and the solution, rather than as a phenomenon for which the membrane is responsible. The membrane is merely the artifice by which this difference is made manifest.
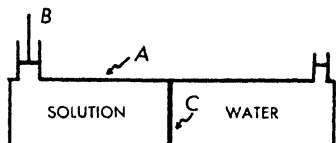
Fig. 6. Osmotic Pressure of Solutions

Measurement of Osmotic Pressure. Pfeffer's original apparatus for measuring the osmotic pressure of solutions consisted of a porous clay cup in the pores of which was deposited a strong film of copper ferrocyanide. To this cup was attached a mercury manometer; the other end of the manometer contained air and was sealed. The cell, after being filled completely with the solution whose osmotic pressure was to be determined, was closed and then immersed in water at a constant temperature. The tendency of solvent to enter the solution developed a pressure within the cell which could be read from the manometer. The difference between the final pressure at equilibrium and the initial pressure was taken as the osmotic pressure of the solution.

Pfeffer's measurements were confined primarily to dilute solutions, and hence low osmotic pressures, and were, at best, not very accurate. Little further progress in this field was forthcoming until H. N. Morse, J. C. W. Frazer, and their collaborators (1901–1923) in America, and the Earl of Berkeley and E. G. J. Hartley (1906–1909) in England, began their celebrated researches on osmotic pressure determination. These men developed methods of preparing osmotic cells in which the membranes were more truly semipermeable and which could withstand much higher pressures than Pfeffer's cells. They also improved the technique of pressure measurement to a point where highly accurate data could

be obtained even with concentrated solutions. Thus, with the apparatus used by Frazer and Morse osmotic pressures could be determined up to 270 atmospheres, while with that of Berkeley and Hartley up to 150 atmospheres.

Only the method of Berkeley and Hartley will be described. A schematic diagram of their apparatus is shown in Fig. 7. A was a porous tube on the outside of which was deposited by a special technique a layer of copper ferrocyanide. This tube was mounted by means of water-tight joints within an outer metal jacket B, which carried an attachment C through which pressure could be applied.
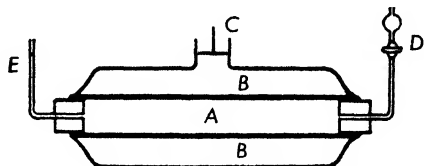
Fig. 7. Osmotic Pressure Apparatus of Berkeley and Hartley

A was filled through D with pure water up to a definite mark on the capillary E, B with the solution under test. The apparatus was then immersed in a thermostat. As a result of osmosis the liquid level in E tended to drop. By applying pressure through C the liquid level at E could be restored to its initial position, and the pressure necessary to accomplish this restoration was taken as the osmotic pressure of the solution. With this method equilibrium is established rapidly, and the concentration of the solution is not changed by dilution with solvent.

TABLE 6

OSMOTIC PRESSURE OF AQUEOUS SUCROSE SOLUTIONS AT 14° C

C (moles/liter of solution)	Π (atm)	Π/C
0.0588	1.34	22.8
0.0809	2.00	24.7
0.1189	2.75	23.2
0.1794	4.04	22.5

Results of Osmotic Pressure Measurements. Some of Pfeffer's data on the osmotic pressures of solutions of cane sugar in water are reproduced in Tables 6 and 7. Table 6 gives the dependence of the osmotic pressure Π on the concentration at constant temperature, while Table 7 gives the variation of osmotic pressure with temperature at constant concentration. These data are reproduced here because of their importance to the theory of solutions, since it was upon these data that van't Hoff based his celebrated analogy between osmotic pressure and gas pressure to be discussed below. The significant facts to be observed about these data are: (1) that at any given temperature the osmotic pressure is directly proportional to concentration, as is indicated by the constancy of the ratio Π/C in the last column of Table 6; and (2) that

for any given concentration the osmotic pressure is directly proportional to the absolute temperature, as is evidenced by the constancy of the ratio Π/T in Table 7.

TABLE 7

OSMOTIC PRESSURE OF AQUEOUS SUCROSE SOLUTIONS AT VARIOUS TEMPERATURES
(1 Per Cent Solution by Weight)

T	Π (atm)	Π/T
273.0	0.649	0.00238
279.8	0.664	0.00237
286.7	0.691	0.00241
288.5	0.684	0.00237
295.0	0.721	0.00244
305.0	0.716	0.00235
309.0	0.746	0.00241

More extended and accurate data on the osmotic pressures of aqueous sucrose solutions at various temperatures are shown in Table 8. They are all based on the measurements of Berkeley and Hartley and of Morse, Frazer, and their co-workers, and are typical of the results obtained with nonelectrolytes in water solutions.

TABLE 8

OSMOTIC PRESSURES OF AQUEOUS SUCROSE SOLUTIONS

m (moles/1000 g H_2O)	Osmotic Pressures (atm)				
	0°	20°	40°	60°	80°
0.1	2.46	2.59	2.56	2.72	
0.2	4.72	5.06	5.16	5.44	
0.3	7.09	7.61	7.84	8.14	
0.4	9.44	10.14	10.60	10.87	
0.5	11.90	12.75	13.36	13.67	
0.6	14.38	15.39	16.15	16.54	
0.7	16.89	18.13	18.93	19.40	
0.8	19.48	20.91	21.80	22.33	23.06
0.9	22.12	23.72	24.74	25.27	25.92
1.0	24.83	26.64	27.70	28.37	28.00

The van't Hoff Equation for Osmotic Pressure. van't Hoff first called attention to the fact that the osmotic pressure of a solution varies with concentration and temperature in exactly the same manner as does the pressure of an ideal gas. Since Pfeffer's data indicated that the osmotic pressure is proportional to concentration and temperature, it must follow that

$$\Pi = kCT \tag{27}$$

But $C = n/V$, where n is the number of moles of solute contained in V liters of solution. Therefore,

$$\Pi = \frac{knT}{V}$$

$$\Pi V = nkT \tag{28}$$

The analogy between equation (28) and the ideal gas law $PV = nRT$ is extremely striking. The analogy becomes even more striking when it is found that the constant k is essentially identical with the gas constant R, as may be seen from the following calculation. The osmotic pressure of a solution containing 45.0 grams of sucrose per liter of solution at $0°$ C is 2.97 atmospheres. Consequently, $\Pi = 2.97$ atmospheres, $V = 1$ liter, $n = 45.0/342.2$, $T = 273.2$, and

$$k = \frac{\Pi V}{nT} = \frac{2.97 \times 1}{\dfrac{45.0}{342.2} \times 273.2} = 0.0827 \text{ liter-atm/degree/mole}$$

This value of k agrees quite closely with $R = 0.0821$ liter-atmosphere per degree per mole, and hence equation (28) may be written as

$$\Pi V = nRT \tag{29}$$

to yield for the osmotic pressure of solutions of nonelectrolytes an equation identical with that for the ideal gas law. The only difference is that equation (29) contains the osmotic pressure Π in the solution instead of the gas pressure P.

According to the van't Hoff equation the osmotic pressure of a solution at any given temperature should depend only on the concentration and not at all on the nature of the solute; i.e., the osmotic pressure should be a colligative property of the solution. This, however, is strictly true only for very dilute solutions, for it is only to such solutions that the van't Hoff equation has been found to be applicable. In more concentrated solutions the observed osmotic pressures are considerably higher than those predicted by equation (29), as may be seen from Table 9. Hence the van't Hoff equation, like the ideal gas law, must be considered to be essentially a limiting law for osmotic pressure rather than one of general validity. This point will be brought out more clearly when the relation of osmotic pressure to vapor pressure lowering is considered.

Various attempts have been made to modify equation (29) in order to bring it into better concordance with measured osmotic pressures. The simplest of these is the observation made by Morse that better results can be obtained when the concentration is expressed in moles

per 1000 grams of solvent, i.e., molality, rather than in moles per liter of solution. As the last column in Table 9 indicates, this manner of calculation improves the agreement between observed and calculated osmotic pressures at the higher concentrations, but the concordance is still far from complete.

TABLE 9

COMPARISON OF OBSERVED AND CALCULATED OSMOTIC PRESSURE OF
SUCROSE SOLUTIONS AT 0° C

Moles/liter of Solution (C)	II Observed (atm)	II Calc. II $= CRT$ (van't Hoff)	Moles/ 1000 g H_2O (m)	II Calc. II $= mRT$ (Morse)
0.02922	0.65	0.655	0.02940	0.659
0.05843	1.27	1.310	0.05914	1.33
0.0970	2.23	2.18	0.0992	2.22
0.1315	2.91	2.95	0.1352	3.03
0.2739	6.23	6.14	0.2905	6.51
0.4406	11.8	9.88	0.4852	10.9
0.5328	14.21	11.95	0.5999	13.45
0.7540	21.87	16.91	0.8943	20.05
0.8183	24.55	18.31	0.9919	23.83
0.8766	26.8	19.7	1.073	24.1

Relation of Osmotic Pressure to Vapor Pressure. From purely thermodynamic considerations it is possible to derive a relation between osmotic pressure and the vapor pressure lowering of a solution. This relation is given by

$$\Pi = \frac{RT}{v} \ln \frac{P^0}{P} \tag{30}$$

where Π is the osmotic pressure, R the gas constant, T the absolute temperature, v the volume of *1 mole* of solvent, P^0 the vapor pressure of pure solvent at temperature T, and P the vapor pressure of solvent above the solution whose osmotic pressure is Π. The only assumptions involved in the derivation of equation (30) are that the solvent vapor behaves ideally and that the molar volumes of pure solvent and solvent in solution are equal. As neither of these assumptions is particularly serious, equation (30) may be expected to reproduce the osmotic pressures up to fairly high concentrations. Table 10 shows that such is actually the case. The difficulty with using equation (30) is the dearth, in most instances, of accurate data on the vapor pressures of solutions.

A word concerning the units of equation (30) is in order. The units in which Π is obtained depend on the units employed for v and R. If v is expressed in liters and R in liter-atmospheres per degree per mole, then

Π will be given in atmospheres. The same result can be obtained for v in cubic centimeters and R in cubic centimeter-atmospheres.

TABLE 10

COMPARISON OF OBSERVED AND CALCULATED OSMOTIC PRESSURES
OF SUCROSE SOLUTIONS

($t = 30°$ C)

Moles of Sucrose per 1000 g H_2O	Π Obs. (atm)	Π Calc. ($\Pi V = nRT$)	Π Calc. (equation 30)
0.1	2.47	2.40	2.44
1.0	27.22	20.4	27.0
2.0	58.37	35.1	58.5
3.0	95.16	45.5	96.2
4.0	138.96	55.7	138.5
5.0	187.3	64.5	183.0
6.0	232.3	—	231.0

The van't Hoff equation $\Pi V = nRT$ can readily be deduced from equation (30) for the special case of dilute solutions obeying Raoult's law. For a solution obeying Raoult's law, $P/P^0 = N_1$, and hence

$$\Pi v = RT \ln \frac{P^0}{P}$$
$$= - RT \ln N_1$$
$$= - RT \ln (1 - N_2)$$

If $\ln (1 - N_2)$ is expanded now in series as before, then for dilute solutions all terms beyond the first can be neglected, and $\ln (1 - N_2)$ becomes $- N_2 = - n/n_1$. Therefore,

$$\Pi v = \frac{RTn}{n_1}$$
$$\Pi(vn_1) = nRT$$

But vn_1 is the total volume of solvent containing n moles of solute, which for dilute solutions is essentially the volume V of the solution. Consequently,

$$\Pi V = nRT$$

which is identical with equation (29). In the light of this derivation of equation (29) from equation (30) it is readily understandable why the van't Hoff equation is limited in applicability only to very dilute solutions and is, in essence, the limiting law for the osmotic pressure of a solution just as $PV = nRT$ is the limiting law for the behavior of any real gas.

Molecular Weight Calculation from Osmotic Pressures. Since the van't Hoff equation is valid for dilute solutions of solutes in a solvent,

it may be employed to calculate the molecular weight of the solutes from osmotic pressure measurements on such solutions. Substituting for n the equivalent W/M, where W is the weight of solute of molecular weight M dissolved in V liters of solution, equation (29) becomes

$$\Pi V = \frac{W}{M} RT \tag{31}$$

and therefore,
$$M = \frac{WRT}{\Pi V} \tag{32}$$

To illustrate the calculation of molecular weights from osmotic pressures, we may take the data of Flusin (1901) on the osmotic pressure of a solution of antipyrine in water. Flusin found that a solution of 10 grams of antipyrine ($C_{11}H_{12}N_2O$) in a liter of solution gave an osmotic pressure of 1.18 atmospheres at 0° C. Inserting these data into (32), we find for the molecular weight:

$$M = \frac{WRT}{\Pi V}$$
$$= 10 \times \frac{0.0821 \times 273.2}{1.18 \times 1}$$
$$= 190$$

The molecular weight estimated from the atomic weights is 188.2.

Although this method may be employed for molecular weight determinations of dissolved solutes, it is rarely used because of the extreme difficulty in obtaining accurate osmotic pressure data for dilute solutions. The same information can be obtained more readily by, say, a freezing point lowering determination with an accuracy that can hardly be attained in osmotic pressure measurements.

Theories of Osmotic Pressure. Various theories have been proposed to explain the nature and cause of osmosis and osmotic pressure. The earliest of these is the solute bombardment theory of van't Hoff. The analogy between the van't Hoff equation for osmotic pressure and the ideal gas law suggested to van't Hoff that the nature and cause of osmotic pressure may be the same as that of gas pressure, namely, molecular bombardment. According to the kinetic theory, gas pressure is believed to result from the impacts between gas molecules distributed in a container and the walls of the container. Similarly van't Hoff considered the solute molecules in a solution to be distributed throughout the solvent, and ascribed the osmotic pressure as due to the collisions of solute molecules with the semipermeable membrane. Assuming the solvent to have no action of any kind, he pointed out that the molecules of solute may be thought of as behaving like a gas of the same concentration con-

fined in a volume equal to that of the solution. Therefore, the osmotic pressure is essentially the pressure which an ideal gas would exert if it were to occupy the same volume as that of the solution. With this picture the identity of the van't Hoff equation and the ideal gas law is understandable. In one case the pressure is the result of bombardment by gas molecules, in the other the result of solute molecules behaving as if they were an ideal gas at the same concentration.

Several objections have been raised against the solute bombardment theory. In this theory the solvent was assumed to play no role and to behave as if the space occupied by the solvent were all space free for the solute molecules to move in. Yet this can hardly be the case. First, as will be shown, the pressure exerted by the solvent on a semipermeable membrane cannot be the same on the solvent and solution sides. Second, the space occupied by solvent molecules can scarcely be considered free space, and hence for the same concentration in the gas phase and in solution the osmotic and gas pressures cannot be the same. For these reasons, and others, the tendency at present is not to regard osmotic pressure as due to bombardment of the semipermeable membrane by solute molecules.

An alternate explanation is offered by the solvent bombardment theory. In this theory the osmosis and osmotic pressure are believed to result from the unequal bombardment of the semipermeable membrane by solvent molecules in the solution and in the pure solvent. On the pure solvent side every molecule that hits the membrane exerts solvent pressure and has a chance of passing through the membrane. On the solution side, however, only part of the collisions with the membrane are due to solvent molecules, the rest being due to solute. Consequently, the pressure exerted by solvent molecules in the solution must be less than the pressure exerted by solvent molecules in the pure solvent, and a tendency develops for solvent to pass into the solution. The difference in the solvent pressures exerted in the pure solvent and in the solution is taken as the osmotic pressure of the solution. Since the difference in solvent pressures developed will be greater the greater the concentration of solute molecules in solution, the osmotic pressure, on this basis, should also be greater the higher the concentration, as is actually observed.

Still a third theory ascribes osmosis as due to the fact that the vapor pressure of the solvent is greater in the pure solvent than in solution. As a result of this difference there is a tendency for solvent to distill through the membrane into the solution until the two pressures become equal. This distillation can be prevented by applying pressure to the solution. Since it can be shown thermodynamically that the vapor pressure of a liquid increases with increase in total applied pressure, applica-

tion of pressure to the solution raises the vapor pressure of the solvent in the latter until it becomes equal to that of the pure solvent. Under these conditions osmosis stops, and the pressure required to accomplish this is taken as the osmotic pressure.

At present there is no way of distinguishing between the solvent bombardment and the vapor pressure theories of osmotic pressure. It may well be that both are merely different aspects of the same phenomenon and are equally valid explanations of osmosis and osmotic pressure.

REFERENCES FOR FURTHER READING

1. J. A. V. Butler, "Fundamentals of Chemical Thermodynamics," The Macmillan Company, New York, 1935, Vols. I and II.
2. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.
3. K. Jellinek, "Lehrbuch der physikalischen Chemie," Ferdinand Enke, Stuttgart, 1928, Vol. II.
4. F. H. MacDougall, "Physical Chemistry," The Macmillan Company, New York, 1936.
5. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931, Vol. I.

PROBLEMS

1. A solution contains 5 g of urea per 100 g of water. What will be the vapor pressure of this solution at 25° C? The vapor pressure of pure H_2O at this temperature is 23.756 mm.   *Ans.* 23.404 mm
2. A solution composed of 10 g of a nonvolatile organic solute in 100 g of diethyl ether has a vapor pressure of 426.0 mm at 20° C. If the vapor pressure of the pure ether is 442.2 mm at the same temperature, what is the molecular weight of the solute?
3. At 25° C, 10.50 liters of pure N_2, measured at atmospheric pressure, are passed through an aqueous solution of a nonvolatile solute, whereby the solution loses 0.2455 g in weight. What is the vapor pressure of the solution and the mole fraction of solute?
4. Select the proper data and calculate the ebullioscopic constant for C_6H_6.
5. From the data in Table 3 calculate the molar heat of vaporization of ethyl alcohol and compare it with the observed value.   *Ans.* 9420 cal/mole
6. If 30 g of diphenyl are dissolved in 250 g of benzene, what will be the boiling point of the resulting solution under atmospheric pressure?   *Ans.* 82.1° C
7. A solution consisting of 5.00 g of an organic solute per 25.00 g of CCl_4 boils at 81.5° C under atmospheric pressure. What is the molecular weight of the solute?
8. Select the necessary data and calculate the cryoscopic constant of cyclohexane.
9. From the data given in Table 5, calculate the heat of fusion per mole of phenol and compare your answer with the observed value.
10. What weight of glycerol would have to be added to 1000 g of water in order to lower its freezing point 10° C?   *Ans.* 495 g
11. An aqueous solution contains 5% by weight of urea and 10% by weight of glucose. What will be its freezing point?
12. Compare the weights of methanol and glycerol which would be required to lower the freezing point of 1000 g of water 1° C.

13. A sample of CH_3COOH is found to freeze at 16.4° C. Assuming that no solid solution is formed, what is the concentration of impurities in the sample?

14. A mixture which contains 0.550 g of camphor and 0.045 g of an organic solute freezes at 157.0° C. The solute contains 93.46% of C and 6.54% by weight of H. What is the molecular formula of the compound? *Ans.* $C_{12}H_{10}$

15. A solution contains 2% of glucose by weight. What is its osmotic pressure at 25° C? *Ans.* 2.72 atm

16. An aqueous solution freezes at $-1.50°$ C. Calculate (a) the normal boiling point, (b) the vapor pressure at 25° C, and (c) the osmotic pressure at 25° C of the given solution.

17. The average osmotic pressure of the human blood is 7.7 atm. What should be the total concentration of various solutes in the blood? What should be the freezing point of blood?

18. The vapor pressure of an aqueous solution at 25° C is 23.45 mm. Calculate its osmotic pressure given that the vapor pressure of pure H_2O is 23.756 mm at 25° C.

Solutions of Electrolytes

Introduction. Solutions of substances in water or other solvents which do not conduct electricity are said to be *nonelectrolytes*. Such solutions exhibit the colligative behavior described in the preceding chapter and obey in dilute solutions the various relations deduced there. These are also the solutions which yield for the dissolved substances normal molecular weights, i.e., the molecular weights expected from their chemical formulas, or, occasionally, some simple multiple of it. On the other hand, there is a class of substances, particularly salts and inorganic acids and bases, which when dissolved in water or other appropriate solvents yields solutions which conduct electricity to a greater or lesser extent. Such solutions are said to be *electrolytes*. Solutions of electrolytes, like those of nonelectrolytes, exhibit the colligative properties of vapor pressure lowering, boiling point elevation, freezing point lowering, and osmotic pressure, but they do not obey the simple relations deduced for nonelectrolytes. The colligative effects observed are always *greater* than those to be expected from the concentration. Stated differently, solutions of electrolytes behave as if the dissolved substance had in solution a molecular weight *lower* than the simplest formula weight of the substance.

This difference in the colligative behavior of electrolytes and nonelectrolytes led some early research workers to ascribe to electrolytes "anomalous" colligative properties. However, it is realized at present that solutions of electrolytes are only "anomalous" in the sense that they do not obey the laws formulated for dilute solutions of nonelectrolytes. When certain effects present in electrolytes but absent in nonelectrolytes are taken into account, a great deal of the "anomaly" disappears, and the behavior of electrolytes becomes more understandable. This chapter will be devoted to a discussion of the colligative properties of electrolytic solutions and to an exposition of some of the theories which have been advanced to explain their "anomaly." Certain other aspects of electrolyte behavior will be elaborated on in later chapters.

Colligative Properties of Electrolytes. As was indicated above, the freezing point lowering, boiling point elevation, vapor pressure low-

ering, and osmotic pressure of solutions of electrolytes all are higher than the corresponding effects for solutions of nonelectrolytes of the same total concentration. The freezing point data shown in Table 1 may be taken as typical of the nature of the deviations. This table lists the ratios of observed freezing point lowering, ΔT_f, to molality, m, at various concentrations for a number of electrolytes in water solution. According to the arguments developed in the preceding chapter, this ratio should approach for dilute aqueous solutions the value of K_f for water, namely, 1.86° per mole per 1000 grams of solvent. Inspection of the table reveals, however, that the limiting values approached by the various electrolytes are considerably higher than 1.86°. Furthermore, the limit approached is not the same throughout, but varies from approximately $2 \times 1.86 = 3.72°$ for substances like hydrochloric acid and ammonium chloride to $3 \times 1.86 = 5.58°$ for cobalt chloride, and $4 \times 1.86 = 7.44°$ for potassium ferricyanide.

TABLE 1

$\Delta T_f/m$ FOR AQUEOUS SOLUTIONS OF ELECTROLYTES

m	HCl	HNO$_3$	NH$_4$Cl	CuSO$_4$	H$_2$SO$_4$	CoCl$_2$	K$_2$SO$_4$	K$_3$Fe(CN)$_6$
0.0005	—	—	—	—	—	—	—	7.3
0.001	3.690	—	—	—	—	—	5.280	7.10
0.002	3.669	—	—	—	—	5.35	—	6.87
0.0025	—	—	—	3.003	5.052	—	5.258	—
0.005	3.635	3.67	3.617	2.871	4.814	5.208	5.150	6.53
0.01	3.601	3.64	3.582	2.703	4.584	5.107	5.010	6.26
0.05	3.532	3.55	3.489	2.266	4.112	4.918	4.559	5.60
0.10	3.523	3.51	3.442	2.08	3.940	4.882	4.319	5.30
0.20	3.54	3.47	3.392	1.91	3.790	4.946	4.044	5.0
0.40	—	3.46	—	—	3.68	5.170	3.79	—
1.00	3.94	3.58	3.33	1.72	4.04	6.31	—	—
2.00	4.43	3.79	3.34	—	5.07	8.51	—	—
4.00	5.65	4.16	3.35	—	7.05	—	—	—

To represent the colligative properties of electrolytes by means of the relations for nonelectrolytes, van't Hoff suggested the use of a factor i, known as the *van't Hoff factor*. The van't Hoff factor is defined as *the ratio of the colligative effect produced by a concentration m of electrolyte divided by the effect observed for the same concentration of nonelectrolyte.* In other words, the van't Hoff factor indicates the number of times the colligative property of an electrolyte is greater than the effect which would be produced by the same concentration of nonelectrolyte. Applying this definition of i to freezing point depression of electrolyte solutions, it follows that

$$i = \frac{\Delta T_f}{(\Delta T_f)_0} \tag{1}$$

where ΔT_f is the freezing point lowering for the electrolyte and $(\Delta T_f)_0$ is the freezing point depression for a nonelectrolyte of the same concentration. Since, according to equation (22) of Chapter VI, $(\Delta T_f)_0 = K_f m$, then

$$i = \frac{\Delta T_f}{K_f m}$$

and therefore, $$\Delta T_f = i K_f m \qquad (2)$$

The values of i must be calculated from experimental data for each electrolyte at various concentrations. It has been found, however, that once i is known for a particular concentration of an electrolyte for one of the colligative properties, say freezing point lowering, the same value of i, within a small temperature correction where necessary, is essentially valid for the other properties at the same concentration. Consequently, we may write

$$i = \frac{\Delta T_f}{(\Delta T_f)_0} = \frac{\Delta T_b}{(\Delta T_b)_0} = \frac{\Delta P}{(\Delta P)_0} = \frac{\Pi}{(\Pi)_0} \qquad (3)$$

where the quantities without subscripts refer to the electrolyte and those with subscripts to the nonelectrolyte of the same concentration. From equation (3) and equations (2), (13), and (29) of the last chapter, the expressions for the boiling point elevation, vapor pressure lowering, and osmotic pressure for solutions of electrolytes become:

$$\Delta T_b = i(\Delta T_b)_0 = i K_b m \qquad (4)$$

$$\Delta P = i(\Delta P)_0 = i P^0 N_2 \qquad (5)$$

$$\Pi = i(\Pi)_0 = \frac{i n R T}{V} \qquad (6)$$

These equations may be expected to be applicable only to dilute solutions.

Table 2 lists the values of i calculated from Table 1 by means of equation (2). A perusal of this table reveals that in dilute solutions i increases as the molality is lowered and approaches a limit of two for electrolytes such as hydrochloric acid, nitric acid, ammonium chloride, and copper sulfate, a limit of three for electrolytes of the type sulfuric acid, cobalt chloride, and potassium sulfate, and a limit of four for potassium ferricyanide. In more concentrated solutions, on the other hand, i passes through a minimum and then increases, the increase frequently rising above the limits established in dilute solutions, as is the case with hydrochloric acid, nitric acid, sulfuric acid, and cobalt chloride. Obviously, such large divergences cannot be ascribed to deviations from the conditions prevailing in solutions of nonelectrolytes, and hence a more fundamental explanation of the nature of electrolytic solutions is necessary. Such an explanation was first supplied by the Arrhenius theory of electrolytic dissociation.

TABLE 2

VAN'T HOFF FACTORS, i, FOR VARIOUS ELECTROLYTES

m	HCl	HNO$_3$	NH$_4$Cl	CuSO$_4$	H$_2$SO$_4$	CoCl$_2$	K$_2$SO$_4$	K$_3$Fe(CN)$_6$
0.0005								3.92
0.001	1.98						2.84	3.82
0.002	1.97					2.88		3.70
0.0025				1.61	2.72		2.83	
0.005	1.95	1.97	1.95	1.54	2.59	2.80	2.77	3.51
0.01	1.94	1.96	1.92	1.45	2.46	2.75	2.70	3.31
0.05	1.90	1.91	1.88	1.22	2.21	2.64	2.45	3.01
0.10	1.89	1.89	1.85	1.12	2.12	2.62	2.32	2.85
0.20	1.90	1.87	1.82	1.03	2.04	2.66	2.17	2.69
0.40		1.86			1.98	2.78	2.04	
1.00	2.12	1.92	1.79	0.93	2.17	3.40		
2.00	2.38	2.04	1.80		2.73	4.58		
4.00	3.04	2.24	1.80		3.79			

The Arrhenius Theory of Electrolytic Dissociation. The "anomalies" encountered in the colligative properties of electrolytes and the fact that solutions of electrolytes conduct electricity led Svante Arrhenius to propose in 1887 his celebrated *theory of electrolytic dissociation*. The essential points of this theory are already familiar to the student from his elementary chemical studies. Arrhenius postulated that electrolytes in solution are dissociated into electrically charged particles, called ions, in such a manner that the total charge on the positive ions is equal to the total charge on the negative ions. The net result is, therefore, that the solution as a whole is neutral in spite of the presence of electrically charged particles in it. Further, the ions formed from an electrolyte may be composed of a single element or a group of elements, radicals, and, again, their charges may differ depending on the nature of the ions. Thus, sodium chloride is considered to be dissociated in Na$^+$ and Cl$^-$, sulfuric acid into 2 H$^+$ and SO$_4^{--}$, barium sulfate into Ba^{++} and SO$_4^{--}$, and potassium ferricyanide into 3 K$^+$ and Fe(CN)$_6^{---}$. The charge on an ion corresponds to its valence and is, in general, positive for the metals and negative for the nonmetals and nonmetallic radicals.

Once the presence of ions in an electrolytic solution is granted, the conductance of electricity through such a solution can readily be explained. If a pair of electrodes is dipped into a solution of an electrolyte and a potential applied across them, the positive ions will be attracted to the negative electrode and will migrate toward it, while the negative ions will be attracted toward the positive electrode and will move in that direction. Such a migration of ions through a solution constitutes

a flow of electricity through the solution. Since nonelectrolytes do not yield ions, no migration of charged particles is possible, and hence solutions of this type are nonconducting.

Arrhenius pointed out further that an electrolyte in solution need not necessarily be completely dissociated into ions; instead it may be only partially dissociated to yield ions in *equilibrium* with unionized molecules of the substance. It may then be anticipated from the laws of chemical equilibrium that the extent of dissociation will vary with concentration, becoming greater as the concentration of dissolved substance becomes lower. In view of this, complete dissociation may be expected to take place only in infinitely dilute solutions. At finite concentrations, however, the electrolyte will be only partially ionized to a degree dependent on the nature of the substance and the concentration.

This idea of partial electrolytic dissociation was employed by Arrhenius to explain the colligative behavior of solutions of electrolytes. The colligative properties of a dilute solution depend on the number of particles, irrespective of kind, present in a given quantity of solvent. When there is no dissociation, dilute solutions exhibit the properties described and formulated in the preceding chapter. When a substance in solution dissociates into ions, however, the number of particles in solution is increased. If it is assumed now that an ion acts with respect to the colligative properties in the same manner that an unionized molecule does, the increase in the number of particles in solution should cause an increase in the colligative effects. Thus, since the molal freezing point depression for a nonelectrolyte in water is 1.86, it may be anticipated that for an electrolyte like hydrochloric acid, which yields on complete dissociation two ions for every molecule, the molal freezing point depression should be twice as great, namely, $2 \times 1.86 = 3.72°$. Similarly, for an electrolyte like potassium sulfate, yielding three ions, the molal freezing point depression should be $3 \times 1.86° = 5.58$, while for potassium ferricyanide, with four ions, $4 \times 1.86 = 7.44°$. These molal lowerings for the types of electrolytes mentioned are the ones to be anticipated on complete dissociation, and hence these are the values to be observed only in extremely dilute solutions. That such is actually the case is evident from an inspection of Table 1. Electrolytes which yield two ions, such as hydrochloric acid, nitric acid, ammonium chloride, and copper sulfate, approach the value of $\Delta T_f/m = 2 \times 1.86 = 3.72°$ as the molality approaches zero; electrolytes which yield three ions, namely, sulfuric acid, cobalt chloride, and potassium sulfate, approach $3 \times 1.86 = 5.58°$ as a limit; while potassium ferricyanide approaches on dilution a value of $\Delta T_f/m = 4 \times 1.86 = 7.44°$. It is seen, therefore, that the observed limiting molal depressions for the various electrolytes are in accord with the predictions of the Arrhenius theory.

The Degree of Dissociation of Electrolytes. From what has been said above it is evident that in terms of this theory any observed values of molal freezing point depressions higher than 1.86 and lower than the limits which would be reached on complete dissociation are to be accounted for by partial dissociation of the electrolyte. If this be the case, it is possible to calculate the degree of ionization of an electrolyte from the observed colligative data or from the values of i calculated from these. To deduce the method of calculation, consider an electrolyte A_xB_y which dissociates into x ions of A, each of charge z_+, and y ions of B, each of charge z_-, according to the equation

$$A_xB_y = xA^{z+} + yB^{z-} \tag{7}$$

If the original molality of the electrolyte is m, and if we let α be the fraction of one mole dissociated into ions, i.e., the *degree of dissociation*, then the number of moles of A_xB_y that dissociate is $m\alpha$, and the number of moles that remain unionized is $m - m\alpha = m(1 - \alpha)$. But for each mole of A_xB_y that dissociates, x moles of positive ions and y moles of negative ions are obtained. Consequently, for $m\alpha$ moles dissociating, $x(m\alpha)$ moles of A^{z+} and $y(m\alpha)$ moles of B^{z-} are obtained. The total number of moles, m_t, of substances of all types present in solution is then

$$\begin{aligned} m_t &= m(1 - \alpha) + x(m\alpha) + y(m\alpha) \\ &= m[1 - \alpha + x\alpha + y\alpha] \\ &= m[1 + \alpha(x + y - 1)] \end{aligned} \tag{8}$$

Designating by ν the *total number of ions yielded by a molecule of the electrolyte*, $\nu = x + y$, and equation (8) becomes

$$m_t = m[1 + \alpha(\nu - 1)] \tag{9}$$

Now, for a total molality m_t the freezing point depression must be given by $\Delta T_f = K_f m_t$, and hence, in terms of equation (9),

$$\Delta T_f = K_f m[1 + \alpha(\nu - 1)] \tag{10}$$

Solving for α, we obtain

$$\begin{aligned} \alpha &= \frac{\left(\dfrac{\Delta T_f}{K_f m} - 1\right)}{(\nu - 1)} \\ &= \frac{\Delta T_f - K_f m}{(\nu - 1)(K_f m)} \end{aligned} \tag{11}$$

A more general relation for α follows by comparing equation (10) with equation (2). Since according to (2) $\Delta T_f = i K_f m$, i must be given by

$$i = 1 + \alpha(\nu - 1) \tag{12}$$

and, therefore, $$\alpha = \frac{i - 1}{\nu - 1} \tag{13}$$

Equation (13) is applicable to any of the colligative properties, and can just as readily be derived from boiling point elevation, vapor pressure lowering, or osmotic pressure considerations. It gives the degree of dissociation of an electrolyte from a knowledge of i and the type of electrolyte in question. Thus, according to Table 2, i for 0.005 m ammonium chloride is 1.95; and, as $\nu = 2$ for this salt,

$$\alpha = \frac{i-1}{\nu-1} = \frac{1.95-1}{2-1}$$
$$= 0.95$$

or, ammonium chloride is 95 per cent dissociated in 0.005 molal solution according to the Arrhenius theory.

Classification of Electrolytes. Calculations of the degree of dissociation for various electrolytes in aqueous solutions show that practically all salts are highly dissociated into ions. The same is true for strong acids, such as nitric, hydrochloric, hydrobromic, hydriodic, sulfuric, and perchloric, and the bases of the alkali and alkaline earth metals. As may be expected from their high degree of dissociation, aqueous solutions of these substances are good conductors of electricity. On the other hand, there are many substances whose aqueous solutions exhibit relatively poor conductivity and whose colligative behavior indicates that they are only slightly dissociated even at fairly low concentrations. Among these are included a large number of organic acids such as acetic, propionic, and benzoic; inorganic acids such as carbonic, hydrosulfuric, hydrocyanic, orthoarsenic, boric, and hypochlorous acids; and bases such as ammonium, zinc, and lead hydroxides. Solutions of substances that show good conductance and which indicate a high degree of dissociation in solution are designated as *strong electrolytes*. In turn, solutions of substances that exhibit only poor conductance and a low degree of dissociation are called *weak electrolytes*.

However, not all electrolytes can be classified clearly as being either strong or weak. There are some electrolytes, such as aqueous solutions of o-chlorbenzoic, o-nitrobenzoic, 3,5-dinitrobenzoic, and cyanoacetic acids, whose behavior indicates that they are intermediate in properties between the strong and weak electrolytes. Still, the number of substances of intermediate strength is not very large, and hence such electrolytes will not be considered further.

It is frequently convenient to subdivide strong electrolytes further according to the charge of the ions produced. An electrolyte that yields two singly charged ions, such as sodium chloride or nitric acid, is called a uni-univalent electrolyte or an electrolyte of the 1-1 type. Again, an electrolyte that yields univalent positive ions and bivalent negative ions, such as potassium sulfate or sulfuric acid, is called a uni-bivalent electro-

lyte, or 1–2 type, while one that yields bivalent positive ions and univalent negative ions, such as barium chloride or magnesium nitrate, is called a bi-univalent electrolyte, or 2–1 type. Similarly copper sulfate is a bi-bivalent electrolyte, or 2–2 type, while potassium ferricyanide is a uni-trivalent electrolyte, or one of 1–3 type. In this method of classification of strong electrolytes the charge of the positive ions is given first followed by the charge of the negative ions.

Criticism of the Arrhenius Theory. Applications of the Arrhenius theory of electrolytic dissociation to the colligative behavior, electrical conductance, and ionic equilibria of *weak electrolytes* have shown that the theory, as outlined by Arrhenius, is essentially satisfactory for these. However, in attempting to apply this theory to *strong electrolytes* so many anomalies and inconsistencies have been encountered that serious questioning arose early as to the validity of some of Arrhenius's postulates in respect to the nature of strong electrolytes.

Arrhenius showed that on the basis of his theory the degree of dissociation of an electrolyte may be determined not only through i from the colligative properties, but also from conductance measurements by methods that will be described in Chapter XV. The values of α obtained by these two methods agree quite well for weak electrolytes, but for strong electrolytes the agreement is not what is to be expected from the accuracy of the measurements. Thus, although for potassium chloride the α's obtained from freezing point data and conductance agree well up to 0.1 N solutions, the α's for potassium hydroxide and for magnesium sulfate show divergences of 3 to 5 per cent in solutions as dilute as 0.005 N. These discrepancies are considerably outside the limits of experimental error. Again, application of the law of mass action to the partial dissociation equilibrium postulated for electrolyte solutions by Arrhenius shows that this law is obeyed quite well by weak electrolytes, but practically not at all by strong electrolytes.

Another factor which militates against the simple dissociation theory when applied to strong electrolytes is the fact that Arrhenius considered ions in solution to behave as neutral molecules do with respect to the colligative properties; i.e., that solutions containing ions behave essentially as ideal solutions containing the same number of neutral molecules. Yet, such an assumption is hardly plausible. Whereas the forces between neutral molecules in fairly dilute solutions may be relatively small, the electrostatic attractions between electrically charged particles in solutions of electrolytes may be anticipated to exercise a significant effect on the motion and distribution of ions. Consequently such solutions can hardly be ideal in behavior, and ions can hardly be expected to act as if they were neutral and ideal molecules. Still, the effect produced by interionic attractions may be quite small in solutions of weak electrolytes

where the number of ions is not large. Such solutions may behave, therefore, in line with Arrhenius's expectations. But, in strong electrolytes, where the number of ions is large, the effect of interionic attractions should be appreciable, and should be the more pronounced the more concentrated the solution and the higher the valence of the ions.

These considerations, and others, point to the conclusion that the Arrhenius theory, although essentially valid for weak electrolytes, does not represent the true situation in solutions of strong electrolytes. There is no question that strong electrolytes exist in aqueous solutions and in certain other solvents as ions of the type postulated. However, the evidence seems to cast serious doubt on the concept of a partial dissociation equilibrium for strong electrolytes, and on the significance of α for such electrolytes as a degree of dissociation. The consensus of opinion seems to be that the α's calculated from colligative effects and conductance data do give a close approximation to the degree of dissociation of weak electrolytes. On the other hand, the belief is that solutions of strong electrolytes are *completely ionized* even at moderate concentrations, and that the α's calculated for strong electrolytes by either of the methods mentioned merely give an indication of the interionic forces of attraction operating in such solutions.

The Complete Ionization of Strong Electrolytes. That interionic attraction rather than partial dissociation is the dominant effect in strong electrolytes at moderate concentrations can be shown by the use of a quantity called the *osmotic coefficient, g,* introduced by Bjerrum [1] in 1907. This quantity, analogous to the van't Hoff factor i, centers attention on deviations from complete dissociation rather than from no dissociation as does i. It is defined as the ratio of an observed colligative property to what the property would be if dissociation were complete, i.e.,

$$g = \frac{\Delta T_f}{\nu(\Delta T_f)_0} = \frac{\Delta T_b}{\nu(\Delta T_b)_0} = \frac{\Delta P}{\nu(\Delta P)_0} = \frac{\Pi}{\nu(\Pi)_0} \qquad (14)$$

where ν, as before, is the number of ions resulting from a molecule of electrolyte. From the definition of i in equation (3) it is evident that, also,

$$g = \frac{i}{\nu} \qquad (15)$$

For completely dissociated electrolytes exhibiting no interionic attraction, $g = 1$. Deviation of g from unity, i.e., $(1 - g)$, is a measure, therefore, of either partial dissociation, or interionic attraction, or both. A study of the variation of $(1 - g)$ with the concentration of ions in solution should give then some indication of the nature of these deviations.

[1] N. Bjerrum, Zeit. für Elektrochemie, **24**, 259 (1907).

Figure 1 shows a plot of $(1 - g)$ against νC, the concentration of ions in solution, for a number of electrolytes of various types. Since at zero ionic concentration dissociation is complete and no interionic attractions are operative, $g = 1$, $(1 - g) = 0$, and all curves start from the origin. These curves not only are applicable to the electrolytes shown in the figure, but are representative also of other electrolytes of the given types. Thus, if the $(1 - g)$ values for sodium chloride were plotted on the same graph, they would coincide at the lower concentrations with those for potassium chloride, those for sulfuric acid would coincide with the potassium sulfate plot, while those for calcium sulfate would fall along the magnesium sulfate curve, etc. In brief, the $(1 - g)$ values for strong electrolytes of a particular ion type when plotted against the ionic concentration all lie on the same curve at the lower concentrations; and individual differences between electrolytes of a given ion type appear only at higher concentrations.
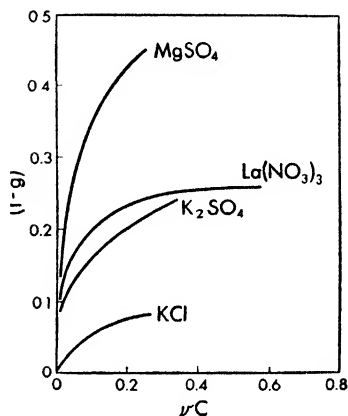
Fig. 1. Plot of $(1 - g)$ vs. νC for Strong Electrolytes

This dependence of the deviations of strong electrolytes in dilute solution on the charges of the ions only is not what may be anticipated from the Arrhenius theory of partial dissociation. This theory makes no allowance for ionic charge, and hence there is no reason why the $(1 - g)$ plots should exhibit the regularity shown. Yet, this is just the type of behavior to be expected if interionic attraction is the determining factor. The electrostatic attraction between charged particles follows the Coulomb law of force, which states that the electrostatic attraction between two charged bodies is directly proportional to their charges and inversely proportional to the square of the distance between them. Since the charges of the ions are determined by their valences, interionic attraction should follow parallel with the valence types of the ions involved, as is the case. We are led to the conclusion, therefore, that *in strong electrolytes*, at least at moderate concentrations, *interionic attraction and not partial dissociation is the important influence;* and that in all probability the deviations exhibited by strong electrolytes are due entirely to electrostatic attractions between ions rather than to incomplete ionization of the electrolyte.

Other evidence which lends considerable credence to the concept of complete ionization of strong electrolytes is supplied by the results of x-ray analysis of crystals. As the student will recall from Chapter IV,

a study of the structure of crystals of potassium and sodium chloride, sodium sulfate, etc., by means of x-ray diffraction indicates that in such crystals the units composing the lattice are ions. There is no entity which corresponds to our ordinary concept of a molecule. In other words, such substances are already *completely ionized* in the solid state. Consequently, when, say, a crystal of sodium chloride is disintegrated by solution in water, the particles breaking away from the crystal and passing into solution are ions and not molecules. If any molecules of sodium chloride are to be present in solution, they must be formed there by *association* of ions rather than by dissociation of molecules already in existence. Since such an association is not very probable at very low concentrations, we again arrive at the deduction that at least in dilute solutions strong electrolytes are completely ionized. And again, if molecules of electrolytes like barium or sodium chloride or potassium sulfate are to be postulated in more concentrated solutions, these must be formed as secondary products through association of ions, and cannot be considered as the initial substances from which ions result by partial dissociation.

The result of these considerations, and others, is that strong electrolytes are believed now to be completely ionized at least up to moderate concentrations, and that in such solutions no unionized molecules are present. What the situation is in concentrated solutions is not so definite. Some of the difficulties and complexities of such solutions will be elaborated on a little later.

The Debye-Hückel Theory of Interionic Attraction. The first successful *quantitative* approach to the problem of interionic attraction in strong electrolytes was made by S. R. Milner.[1] However, Milner's treatment was so involved mathematically that his work attracted little attention among chemists. It was not until 1923, when P. Debye and E. Hückel [2] first published an alternate and much simpler treatment of the subject, that the theory of interionic attraction came to the fore. Since then this theory has grown to occupy a dominant position in all considerations involving electrolytes and their kinetic and thermodynamic behavior.[3]

At this time attention will be confined to the presentation of the salient qualitative details of the theory, without any attempt at mathematical formulation. Like Arrhenius, Debye and Hückel postulate that strong electrolytes exist in solution as ions of the types mentioned, but, unlike

[1] S. R. Milner, Phil. Mag., **23**, 551 (1912); **25**, 742 (1913); **35**, 214, 352 (1918); Trans. Farad. Soc., **15**, 148 (1919).

[2] Debye and Hückel, Physikalische Zeitschrift, **24**, 185 (1923).

[3] Although this theory will be discussed briefly at this point and will be elaborated on at several other stages in the book, no complete and rigid exposition is possible in an elementary text. Any student interested in further and more complete details is advised to read some of the references mentioned at the end of the chapter.

Arrhenius, these authors do not believe that the anomalous properties of strong electrolytes are due to partial dissociation. Rather, they believe that strong electrolytes, at least in dilute solutions, are completely ionized, and that the effects observed are due to the unequal distribution of ions resulting from interionic attraction. By a mathematical argument these authors showed that because of electrostatic attractions between charged ions, each positive ion in solution must be surrounded on an average with more negative ions than ions of like charge; and conversely, each negative ion must be surrounded on an average with more positive than negative ions. In other words, each ion in solution is surrounded by an *ionic atmosphere whose net charge is opposite to that of the central ion.* They showed, further, that the properties of the electrolyte are determined by the interaction of the central ion with its atmosphere. Since the nature of the atmosphere is determined by the valences of the ions in solution, their concentration, the temperature, and the dielectric constant of the medium, it must follow that these are also the factors controlling the thermodynamic properties of the electrolyte. At any given temperature and in any given solvent, the temperature and dielectric constant are fixed, and hence the properties of the electrolyte should depend only on the charges of the ions and their concentration, and not at all on the specific nature of each electrolyte. These conclusions are strictly valid only in very dilute solutions. This limitation arises from the fact that Debye and Hückel were forced to make certain mathematical simplifications which in essence reduce the applicability of their equations to such solutions.

In the Debye-Hückel theory of interionic attraction the effect of the concentration of the ions in solution does not enter directly, but through a quantity called the *ionic strength* of the solution. This quantity, first introduced by G. N. Lewis even before the advent of the Debye-Hückel theory, is a measure of the electrical environment of the solution, and plays in this theory a role analogous to concentration in the Arrhenius theory. It is defined as

$$\mu = \frac{1}{2} \left(C_1 z_1^2 + C_2 z_2^2 + C_3 z_3^2 + \cdots \right)$$

$$= \frac{1}{2} \Sigma C_i z_i^2 \tag{16}$$

where μ is the ionic strength of the solution, C_1, C_2, C_3, $\cdots$ the concentrations of the various ions in gram ionic weights per liter, while z_1, z_2, z_3 $\cdots$ are the valences of the respective ions. The ionic strength of a solution is equal to the molarity only in the case of 1–1 electrolytes. In all other instances the two are not equal. This may be illustrated with the calculation of the ionic strength of C molar solutions of potassium

chloride, magnesium sulfate, barium chloride, and lanthanum sulfate. In potassium chloride, $C_+ = C_- = C$, $z_+ = z_- = 1$, and therefore

$$\mu_{KCl} = \frac{1}{2}[C(1)^2 + C(1)^2]$$
$$= \frac{2\,C}{2} = C$$

For magnesium sulfate, $C_+ = C_- = C$, $z_+ = z_- = 2$, and

$$\mu_{MgSO_4} = \frac{1}{2}[C(2)^2 + C(2)^2]$$
$$= \frac{8\,C}{2} = 4\,C$$

For barium chloride, $C_+ = C$, $C_- = 2\,C$, $z_+ = 2$, $z_- = 1$. Hence,

$$\mu_{BaCl_2} = \frac{1}{2}[C(2)^2 + 2\,C(1)^2]$$
$$= \frac{6\,C}{2} = 3\,C$$

Finally, for lanthanum sulfate, $C_+ = 2\,C$, $C_- = 3\,C$, $z_+ = 3$, $z_- = 2$, and therefore

$$\mu = \frac{1}{2}[2\,C(3)^2 + 3\,C(2)^2]$$
$$= \frac{30\,C}{2} = 15\,C$$

It is readily apparent from these examples that the ionic strength of a solution is determined not only by the stoichiometric concentration of the electrolyte, but also by the valences of its ions. It should also be emphasized that in calculating the ionic strength of a solution the summation called for in equation (16) not only must be extended to the electrolyte of interest, but must include any and all ionic species present, no matter what their source. Thus the ionic strength of a solution containing $C = 0.10$ potassium chloride in presence of $C = 0.01$ barium chloride is

$$\mu = \frac{1}{2}[0.1(1)^2 + 0.1(1)^2 + 0.01(2)^2 + 2(0.01)(1)^2]$$
$$= 0.13$$

Various indications of the ability of the Debye-Hückel theory to explain the behavior of dilute solutions of strong electrolytes will be given at various appropriate stages in this book. At present it is sufficient to

point out the superiority of this theory over that of Arrhenius in explaining the course of the $(1 - g)$ vs. ionic concentration curves shown in Fig. 1. The Debye-Hückel theory predicts that $(1 - g)$ for very dilute solutions of strong electrolytes in water at $0°$ C should be given by

$$(1 - g) = 0.264(z_+z_-)^{3/2} \sqrt{\nu C} \tag{17}$$

According to this equation, $(1 - g)$ should vary linearly with $\sqrt{\nu C}$ with slopes varying according to the valence type of the electrolyte in question. The extent to which these predictions are verified may be judged from Fig. 2, where a plot of $(1 - g)$ vs. $\sqrt{\nu C}$ is given. The solid lines are the plots of equation (17) for the various ionic types, while the dotted lines show the course of the experimentally determined values. The significant facts to be observed in this comparison are (a) that the Debye-Hückel theory demands that the $(1 - g)$ curves follow the valence type of the electrolyte, as is actually the case, and (b) that the experimentally observed data approach the theoretically predicted curves at the low

Fig. 2. Plot of $(1 - g)$ vs. $\sqrt{\nu C}$ for Various Electrolytes

concentrations. These facts are strong confirmation of the essential validity of the concept of interionic attraction as the cause of the deviations in strong electrolytes.

In conclusion, the present status of the problem of electrolytes may be summarized as follows. For weak electrolytes the Arrhenius theory of partial dissociation is, within minor corrections, adequate. On the other hand, for strong electrolytes this theory is not satisfactory. In such solutions interionic attraction is the dominant factor and affords a better explanation of the behavior of such solutions. On the quantitative and theoretical side the Debye-Hückel theory supplies the explanation for the thermodynamic properties of very dilute and moderately dilute solutions, but this theory has not been extended as yet to concentrated solutions. It is quite probable that for concentrated solutions the treatment presented by the Debye-Hückel theory is considerably oversimplified. In order to understand the properties of concentrated solutions it may be necessary to consider not only interionic attraction, but also such phenomena as interaction between ions resulting in association, interaction of the solvent with ions, and the change in the nature of the solvent as a result of the presence of charged particles. Although the possible importance of these factors is appreciated, and although some

of them have been investigated both theoretically and experimentally, no complete analysis of the complex problems involved here has as yet been made.

REFERENCES FOR FURTHER READING

1. J. A. V. Butler, "Fundamentals of Chemical Thermodynamics," Macmillan & Company, Ltd., London, 1935, Vol. II.
2. C. W. Davies, "The Conductivity of Solutions," Chapman and Hall, Ltd., London, 1933.
3. M. Dole, "Experimental and Theoretical Electrochemistry," McGraw-Hill Book Company, Inc., New York, 1935.
4. H. Falkenhagen, "Electrolytes," Oxford University Press, New York, 1934.
5. K. Jellinek, "Lehrbuch der physikalischen Chemie," Ferdinand Enke, Stuttgart, 1930, Vol. III.
6. V. K. La Mer, Trans. Am. Electrochem. Soc., **41**, 507 (1927).
7. G. Scatchard, Chem. Rev., **13**, 7 (1933).
8. Shedlovsky, Brown, MacInnes, Trans. Am. Electrochem. Soc., **66**, 237 (1934).
9. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931, Vol. I.

PROBLEMS

1. A 0.2 molal aqueous solution of KCl freezes at $-0.680°$. Calculate i and the osmotic pressure at $0°$ C. Assume volume to be that of pure H_2O.

    Ans. $i = 1.83$; $\Pi = 8.2$ atm

2. A 0.4 molal aqueous solution of K_2SO_4 freezes at $-1.52°$ C. Assuming that i is constant with temperature, calculate the vapor pressure at $25°$ C and the normal boiling point of the solution.

3. From the data listed in Table 2, calculate the freezing point and the vapor pressure at $25°$ C of a 2.00 molal $CoCl_2$ solution.

4. A solution of HCl, 6.80% by weight, freezes at $-0.706°$ C. Calculate the apparent molarity and the apparent molecular weight of the HCl.

5. A 0.01 molal solution of $K_3Fe(CN)_6$ freezes at $-0.062°$ C. What is the apparent percentage of dissociation?

6. A 2.00 molal HCl solution freezes at $-4.43°$ C. Calculate the apparent percentage of dissociation and explain your answer.

7. A 0.1 molal solution of a weak electrolyte ionizing into two ions freezes at $-0.208°$ C. Calculate the degree of dissociation.   *Ans.* 0.118

8. At $25°$ C a 0.1 molal solution of CH_3COOH is 1.35% dissociated. Calculate the freezing point and osmotic pressure of the solution. Compare your results with those which you would obtain if the acid did not dissociate.

9. What is the osmotic coefficient of the $K_3Fe(CN)_6$ solution given in problem 5?

    Ans. 0.833

10. At $25°$ C the vapor pressure of a 2.00 molal $CoCl_2$ solution is 20.00 mm Hg. Calculate the osmotic coefficient.

11. Given the freezing points of the following $CoCl_2$ solutions construct a plot of $1 - g$ against $\sqrt{vm}$. Determine from this plot the limiting slope and compare it with that predicted by the Debye-Hückel theory. Assume that $m = C$ for this purpose.

Molality	Freezing Point, ° C
0.002	− 0.0107
0.005	− 0.0261
0.01	− 0.0511
0.05	− 0.246
0.10	− 0.488
0.20	− 0.990

12. Compare the ionic strengths of 0.1 N solutions of HCl, $SrCl_2$, $AlCl_3$, $ZnSO_4$, and $Fe_2(SO_4)_3$.

13. A solution is 0.5 molar in $MgSO_4$, 0.1 molar in $AlCl_3$, and 0.2 molar in $(NH_4)_2SO_4$. What is the total ionic strength?   *Ans.* 3.2

14. Using the data of Table 2, plot on the same graph $1 - g$ against $\sqrt{\mu}$ for H_2SO_4 and HCl solutions for molalities less than 0.1, and determine the limiting slopes of the curves obtained. What is the relation between them?

Surface Phenomena and Colloids

Although under ordinary conditions a solute can be dissolved in a solvent to yield a homogeneous solution containing the solute in molecular or ionic dimensions, under special conditions the solute can be distributed through the solvent as particles of considerably larger size. Such a dispersion does not constitute a true solution but rather a stage of subdivision intermediate between true solution and bulk solute. Substances existing in this intermediate state are said to be in *colloidal dispersion or suspension*. One of the significant consequences attending the colloidal dispersion of a substance is a tremendous increase in free surface. Because of this surface development, particles in colloidal dispersion acquire the ability to attract and retain on their surface dissolved substances and solvent molecules which endow the colloidal particles with many of their observed properties. In view of the importance of the phenomenon of surface interaction in any discussion of colloidal behavior, the experimental and theoretical aspects of adsorption in noncolloidal systems will be outlined before elaborating on the subject of colloids.

ADSORPTION

Definitions. Attention has already been directed to the fact that the molecular forces at the surface of a liquid are in a state of unbalance or unsaturation. The same is true of the surface of a solid, where the molecules or ions in the surface of a crystal do not have all their forces satisfied by union with other particles. As a result of this unsaturation, solid and liquid surfaces tend to satisfy their residual forces by attracting onto and retaining on their surfaces gases or dissolved substances with which they come in contact. This phenomenon of concentration of a substance *on* the surface of a solid or liquid is called *adsorption*. The substance thus attracted to a surface is said to be the *adsorbed phase*, while the substance to which it is attached is the *absorbent*.

*Ad*sorption should be carefully distinguished from *ab*sorption. In the latter process a substance is not only retained on the surface, but passes through the surface to become distributed throughout the body of a

solid or liquid. Thus water is *ab*sorbed by a sponge, or water vapor is *ab*sorbed by anhydrous calcium chloride to form a hydrate; but acetic acid in solution and various gases are *ad*sorbed by charcoal. Where doubt exists as to whether a process is true adsorption or absorption, the noncommittal term *sorption* is sometimes employed.

Adsorption of Gases by Solids. Although it is probable that all solids adsorb gases to some extent, adsorption as a rule is not very pronounced unless an adsorbent possesses a large surface for a given mass. For this reason silica gel and charcoals obtained from various sources, such as wood, bone, coconut shells, and lignite, are particularly effective as adsorbing agents. These substances have a very porous structure, and with their large exposed surfaces can take up appreciable volumes of various gases. The extent of adsorption can further be increased by "activating" the adsorbents in various ways. Thus wood charcoal can be "activated" by heating between 350 and 1000° C in a vacuum or in air, steam, and certain other gases to a point where the adsorption of carbon tetrachloride at 24° C can be increased from 0.011 gram per gram of charcoal to 1.48 grams. The activation involves apparently a distilling out of hydrocarbon impurities from a charcoal, and leads thereby to an exposure of a larger free surface for possible adsorption.

TABLE 1

ADSORPTION OF GASES BY CHARCOAL AT 15° C
(1 g of adsorbent)

Gas	Volume Adsorbed (cc)	Critical Temperature (° K)
H_2	4.7	33
N_2	8.0	126
CO	9.3	134
CH_4	16.2	190
CO_2	48	304
HCl	72	324
H_2S	99	373
NH_3	181	406
Cl_2	235	417
SO_2	380	430

The amount of gas adsorbed by a solid depends on the natures of the adsorbent and gas being adsorbed, the area of the adsorbent, the temperature, and the pressure of the gas. The specificity with which certain gases are adsorbed by a given solid can be seen from Table 1, where the volumes of various gases adsorbed by a gram of charcoal at 15° C are given. The volumes of gas have all been reduced to 0° C and 1 atmosphere pressure. A comparison of the relative volumes of various gases adsorbed by a solid reveals that in general the extent of adsorption parallels the increase in the critical temperature of the gases. This parallel-

ism suggests that gases which liquefy easily are more readily adsorbed, but it does not necessarily indicate that the gases exist as liquids on the surface. A similar correlation is obtained with boiling points.

As may be expected, an increase in the surface area of adsorbent increases the total amount of gas adsorbed. Since the surface area of adsorbing agents cannot be determined readily, the usual practice is to employ the mass of adsorbent as a measure of the surface available, and to express the amount of adsorption per unit mass of adsorbing agent used.

Experimental evidence has definitely shown that in adsorption a true equilibrium is established between the gas in contact with a solid and the gas on the surface; i.e., for a given gas and adsorbent the extent of adsorption under any condition of temperature and pressure is definite and reproducible. Like all equilibria, the adsorption process is strongly affected by temperature. Invariably an increase in temperature leads to a decrease in the amount adsorbed, and vice versa, as may be seen from the following data. At 0° C and 600 mm pressure a gram of charcoal adsorbs about 10 cc of nitrogen. As the temperature is lowered to − 29° C the amount adsorbed at 600 mm increases to 20 cc, and at − 78° C it is 45 cc under the same pressure. This indicates that adsorption of a gas by a solid is accompanied by an evolution of heat. These

Fig. 1. Typical Adsorption Isotherm

heats of adsorption, which have actually been measured in a number of cases, will be referred to a little later.

The manner in which the amount of gas adsorbed by a solid varies with pressure at constant temperature is exemplified by Fig. 1. Such a plot, called an *adsorption isotherm*, is typical of general adsorption behavior. At the lower pressures the amount adsorbed increases relatively rapidly with pressure, and then much more slowly as the surface becomes covered with gas molecules. To represent the variation of the amount of adsorption per unit area or unit mass with pressure, Freundlich proposed on purely empirical grounds the equation

$$y = kP^{\frac{1}{n}} \tag{1}$$

Here y is the weight or volume of gas adsorbed per unit area or unit mass of adsorbent, P is the equilibrium pressure, and k and n are empirical constants dependent on the nature of solid and gas and on the temperature. This equation may be tested as follows. Taking logarithms of both sides, equation (1) becomes

$$\log_{10} y = \log_{10} k + \frac{1}{n} \log_{10} P \qquad (2)$$

If $\log_{10} y$ is plotted now against $\log_{10} P$, a straight line should result with slope equal to $1/n$ and ordinate intercept equal to $\log_{10} k$. Figure 2 shows such a plot constructed from the data given in Table 2 for the adsorption of nitrogen on mica at 90° K. Although the requirements of the equation are met satisfactorily at the lower pressures, at higher pressures the
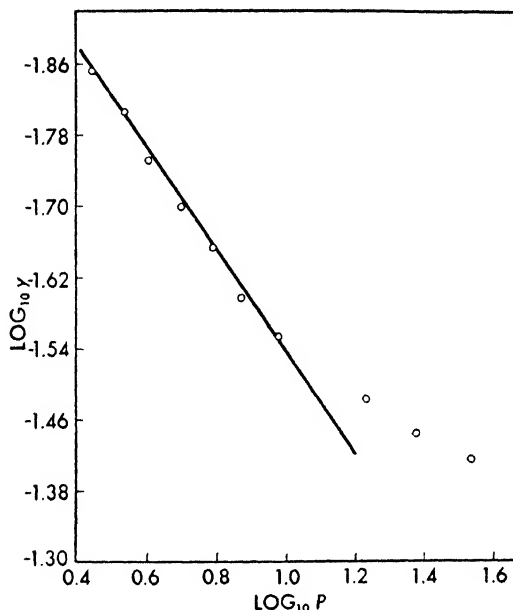
Fig. 2. Application of Freundlich Equation to Adsorption of N_2 on Mica at 90° K

experimental points curve away from the straight line, indicating that this equation does not have general applicability in reproducing adsorption of gases by solids.

The Langmuir Adsorption Equation. A much better adsorption equation was deduced by Irving Langmuir[1] from theoretical considerations. Langmuir first of all postulates that gases in being adsorbed by a solid surface cannot form a layer more than a *single molecule* in depth. Further he visualizes the adsorption process as consisting of two opposing actions, a condensation of molecules from the gas phase onto the surface and an evaporation of molecules from the surface back into the body of the gas. When adsorption first starts, every molecule colliding with the surface may condense on it. However, as adsorption proceeds,

[1] Langmuir, J. Am. Chem. Soc., **38**, 2221 (1916); **40**, 1361 (1918).

only those molecules may be expected to be adsorbed which strike a part of the surface not already covered by adsorbed molecules. The result is that the initial *rate* of condensation of molecules on a surface is highest, and falls off as the area of surface available for adsorption is decreased.

<div align="center">

TABLE 2

ADSORPTION OF NITROGEN ON MICA AT 90° K
(I. Langmuir)

</div>

P in Bars (dynes/square centimeter)	y (mg/square centimeter)
2.8	14.0×10^{-3}
3.4	15.6
4.0	17.6
5.0	19.8
6.1	22.1
7.4	25.2
9.5	27.8
17.0	32.8
23.8	35.8
34.0	38.4

On the other hand, a molecule adsorbed on a surface may, by thermal agitation, become detached from the surface and escape into the gas. The rate at which such desorption will occur will depend, in turn, on the amount of surface covered by molecules, and will increase as the surface becomes more fully saturated. These two rates, condensation and desorption, will eventually become equal, and when this happens an adsorption equilibrium will be established.

These ideas can be formulated mathematically. If we let θ be the *fraction* of the total surface covered by adsorbed molecules at any instant, then the fraction of surface bare and available for adsorption is $(1 - \theta)$. Since, according to kinetic theory, the rate at which molecules strike unit area of a surface is proportional to the pressure of the gas, the rate of condensation of molecules should be determined both by the pressure and the fraction of surface bare, or,

$$\text{Rate of condensation} = k_1(1 - \theta)P$$

where k_1 is a constant of proportionality. On the other hand, if we let k_2 be the rate at which molecules evaporate from unit surface when the surface is fully covered, then for a fraction θ of surface covered the rate of evaporation will be

$$\text{Rate of evaporation} = k_2\theta$$

For adsorption equilibrium these rates must be equal. Therefore,

$$k_1(1 - \theta)P = k_2\theta$$
$$k_2\theta + k_1\theta P = k_1 P$$
$$\theta = \frac{k_1 P}{k_2 + k_1 P}$$
$$= \frac{bP}{1 + bP} \tag{3}$$

where $b = k_1/k_2$. Now, the amount of gas adsorbed per unit area or per unit mass of adsorbent, y, must obviously be proportional to the fraction of surface covered, and hence,

$$y = k\theta = \frac{kbP}{1 + bP}$$
$$= \frac{aP}{1 + bP} \tag{4}$$

where the constant a has been written for the product kb.

Equation (4) is the *Langmuir adsorption isotherm.* The constants a and b are characteristic of the system under consideration and are evaluated from experimental data. Their magnitude depends also on the temperature. At any one temperature the validity of the Langmuir adsorption equation can be verified most conveniently by first dividing both sides of equation (4) by P and then taking reciprocals. The result is

$$\frac{y}{P} = \frac{a}{1 + bP}$$
$$\frac{P}{y} = \frac{1}{a} + \left(\frac{b}{a}\right)P \tag{5}$$

Since a and b are constants, a plot of P/y vs. P in equation (5) should yield a straight line with slope equal to b/a and an ordinate intercept equal to $1/a$. In Fig. 3 is given such a plot based on the nitrogen data listed in Table 2. This excellent straight line definitely confirms the Langmuir adsorption equation, and indicates that it is possible to represent adsorption behavior at constant temperature more satisfactorily with this equation than with the Freundlich isotherm. Furthermore, this test, and many similar ones, lend support to the correctness of Langmuir's mechanism of the adsorption process and to his assumption that the adsorbent is covered only by a unimolecular layer of gas molecules.

However, it should be mentioned that at high pressures, and particularly with readily condensible gases, the Langmuir isotherm is not obeyed so well. Under these conditions it is quite probable that adsorbed layers thicker than one molecule result, and it is also possible that the high pressures lead to liquefaction of gas in the pores of the adsorbent.
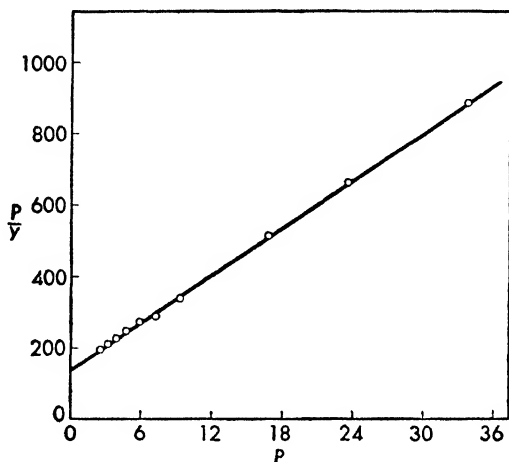
Fig. 3. Application of Langmuir Equation to Adsorption of N_2 on Mica at 90° K

From a graph such as Fig. 3 the constants in equations (4) or (5) are readily determined by taking the slope and the y intercept. Thus from this plot we find that the slope is $b/a = 22.0$, while the y intercept $1/a = 140$. Hence $a = 0.00714$, $b = 0.157$, and the adsorption of nitrogen on mica at 90° K can be represented by the formula

$$y = \frac{0.00714\,P}{1 + 0.157\,P} \tag{6}$$

Types of Adsorption. Careful study of adsorption of various gases on solid surfaces has revealed that the forces operative in adsorption are not the same in all cases. Two types of adsorption are generally recognized: (1) physical or van der Waals adsorption, and (2) chemical or activated adsorption. Physical adsorption is characterized by low heats of adsorption, of the order of 10,000 calories or less, and by the fact that the adsorption equilibrium is reversible and is established rapidly. This is the type observed in the adsorption of various gases on charcoal. The forces responsible for the adsorption here are believed to be of the same kind as are involved in deviations of gases from ideal behavior and in liquefaction, namely, van der Waals forces. On the other hand, activated or chemical adsorption is accompanied by much higher heat changes, ranging from 20,000 to as high as 100,000 calories, and leads to a much

firmer attachment of the gas to the surface. As these heats are of the same order of magnitude as those involved in chemical reactions, it is quite probable that the adsorption consists of a combination of gas molecules with surface atoms to form a surface compound. In the adsorption of oxygen on tungsten it has actually been found that tungsten trioxide distils from the surface at about 1200° K; however, even above this temperature oxygen remains on the surface apparently as WO. As other examples of chemical adsorption may be mentioned that of carbon monoxide on tungsten, oxygen on silver, gold, platinum, and carbon, and hydrogen on nickel.

Undoubtedly many cases of adsorption are neither one type nor the other, but a combination of both. Again, whether physical or chemical adsorption is observed depends sometimes on the temperature. Instances have been observed where physical adsorption at low temperatures passes into chemical adsorption as the temperature is raised. This is true with the adsorption of hydrogen on nickel. In general chemical adsorption is more specific in nature than physical, and is to be found only where there is tendency toward compound formation between a gas and adsorbent. However, since van der Waals forces are not specific in nature, physical adsorption may be found in all instances, although it may possibly be masked by the stronger chemical type.

Adsorption of Solutes by Solids. Solid surfaces can adsorb not only gases, but dissolved substances from a solution as well. When a solution of acetic acid in water is shaken with activated carbon, it is found that part of the acid is removed by the carbon and the concentration of the solution is decreased. Similarly, activated carbon can be used to remove ammonia from solutions of ammonium hydroxide, phenolphthalein from solutions of acids or bases, etc. Adsorption phenomena of this type are encountered also in quantitative analysis. Study has shown that freshly precipitated silver chloride tends to adsorb either silver ions or chloride ions, depending on which is in excess, while arsenic trisulfide tends to adsorb sulfide ions from solutions in which it is precipitated.

As a rule activated carbon is much more effective in adsorbing nonelectrolytes from a solution than electrolytes, and the extent of adsorption is usually the greater the higher the molecular weight of the substance being adsorbed. Conversely, inorganic solids tend to adsorb ions more readily than nonelectrolytes. This tendency of certain adsorbents to attract certain substances in preference to others occasionally leads to the phenomenon of *negative adsorption*, i.e., the concentration of a solute is actually increased after treatment with the adsorbing agent. Dilute solutions of potassium chloride agitated with blood charcoal offer a case in point. The explanation suggested for negative adsorption is

that the solvent, in this case water, is adsorbed in preference to the electrolyte, and as a consequence the concentration of the solute is raised. It is of interest to observe that at high concentrations potassium chloride exhibits positive adsorption, with the salt being adsorbed rather than the water.

Adsorption from solution follows generally the principles laid down for adsorption of gases, and is subject to the same factors. As was pointed out, some adsorbents are specifically more effective in attracting certain substances to their surface than others. Again, an increase in temperature operates to decrease the extent of adsorption, while an increase in surface area increases it. Adsorption of solutes, like that of gases, involves the establishment of an equilibrium between the amount adsorbed on the surface and the concentration of the substance in solution. The variation of extent of adsorption with concentration of solute is usually represented by the Freundlich equation, which seems to work better with this type of adsorption than with gases. For this purpose, equation (1) is written in the form

$$y = kC^{\frac{1}{n}} \tag{7}$$

where y is the mass of substance adsorbed per unit mass of adsorbent, while C is the equilibrium concentration of the solute being adsorbed. k and n are again empirical constants. By taking as before logarithms of both sides of equation (7), we obtain

$$\log_{10} y = \log_{10} k + \frac{1}{n} \log_{10} C \tag{8}$$

and hence a plot of $\log_{10} y$ vs. $\log_{10} C$ should be linear with slope equal to $1/n$ and intercept equal to $\log_{10} k$.

Table 3 gives some data determined by Freundlich for the adsorption of acetic acid on blood charcoal at 25° C, as well as the logarithms of y

TABLE 3

ADSORPTION OF ACETIC ACID FROM WATER SOLUTIONS BY
BLOOD CHARCOAL AT 25° C

(H. Freundlich)

C (moles/liter)	y (g/g charcoal)	$\log_{10} C$	$\log_{10} y$
0.018	28.2×10^{-3}	-1.745	-1.550
0.031	37.2	-1.509	-1.429
0.062	48.0	-1.208	-1.319
0.126	66.5	-0.900	-1.197
0.268	93.0	-0.572	-1.032
0.471	122.	-0.327	-0.914
0.882	149.	-0.055	-0.827

and C. Plotting now $\log_{10} y$ against $\log_{10} C$ the satisfactory straight line shown in Fig. 4 is obtained. From the slope and y intercept of the line we find $n = 2.32$, $k = 0.160$, and hence these data may be represented at $25°$ C by the equation $y = 0.160 \, C^{1/2.32}$.
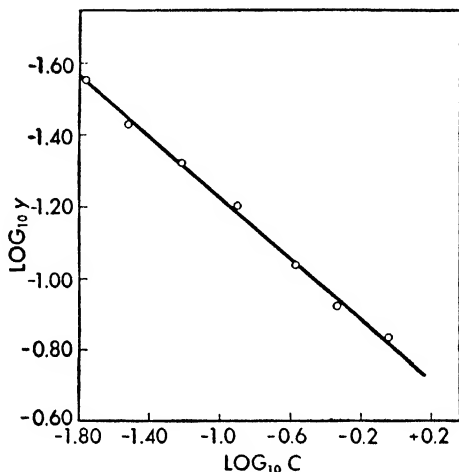
Fig. 4. Application of Freundlich Equation to Adsorption of Acetic Acid on Charcoal

Adsorption at Solution Surfaces. The reader is familiar with the fact that addition of soap to water lowers greatly the surface tension of the latter. Substances which have the ability to produce an appreciable reduction in the surface tension of a liquid on being dissolved in it are called *surface active agents*. In this category are included the soaps, certain sulfonic acids, and various other types of organic compounds. On the other hand, some solutes, and particularly electrolytes, have the faculty of increasing the surface tension of water, and these substances may be said to exhibit negative surface activity.

As early as 1878, J. Willard Gibbs predicted that surface activity is due to unequal distribution of a solute between the surface and body of a solution. By a purely thermodynamic argument he showed that if a solute so distributes itself that unit area of the surface contains q moles of solute in excess of that present in the body of the solution, then for dilute solutions q at equilibrium should be given by

$$q = - \frac{C}{RT} \frac{d\gamma}{dC} \tag{9}$$

where C is the concentration of the solution, T the absolute temperature, R the gas constant, and $d\gamma/dC$ the rate of variation of the surface tension of the solution with concentration. When $d\gamma/dC$ is positive, i.e., when the surface tension of a solution increases with concentration, q

must be negative, and the body of the solution is richer in the solute than the surface, as is the case with many electrolytes. However, when the surface tension of a solution decreases with concentration, $d\gamma/dC$ is negative, q is positive, and the surface contains a higher concentration of solute than the solution. The latter is the case with surface active agents. Positive surface activity, therefore, is associated with what may be considered to be an adsorption of solute from solution by the surface of the solution, while negative surface activity is due to an expulsion of solute from the surface.

Although it is difficult to determine the surface concentration of a solution, equation (9) has been subjected to direct test, particularly through the ingenious techniques developed by McBain and his coworkers.[1] The results obtained on solutions of phenol in water, which show positive surface excess, and on sodium chloride in water, with negative surface adsorption, confirm in an excellent manner the demands of the Gibbs adsorption equation.

Applications of Adsorption. Adsorption finds very extensive application both in the research laboratory and in industry. Adsorption of gases on solids is utilized in the laboratory to preserve the vacuum between walls of Dewar containers designed for the storage of liquid air or liquid hydrogen. Activated charcoal is placed between the walls, and this tends to adsorb any gases which appear due to glass imperfections or diffusion through the glass. Again, adsorption plays a very important part in various aspects of catalysis of gaseous reactions by solid surfaces. This subject will be developed in more detail in a later chapter. Further, all gas masks are merely devices containing an adsorbent or a series of adsorbents which, by preferentially removing poisonous gases from the atmosphere, purify the air for breathing. In a similar manner various adsorbents are used in industry to recover solvent vapors from air, or particular solvents from mixtures of other gases.

As applications of adsorption from solution may be mentioned the clarification of sugar liquors by charcoal, the removal of coloring matter from various other types of solutions, and the recovery of dyes from dilute solutions in a number of solvents. Adsorption has also been used for the recovery and concentration of vitamins and other biologic substances, and finds utility now in a method of analysis called *chromatographic analysis*. This scheme of estimation of small quantities of substance depends on progressive and selective adsorption of a number of constituents present simultaneously in a solution.

Surface active agents find extensive application in detergents, paints, waterproofing, lubrication, and a host of other fields. All surfaces are covered with layers of either gaseous, liquid, or solid films, and the prob-

[1] McBain and Swain, Proc. Roy. Soc. (London), **154A**, 608 (1936).

lem of displacing these frequently becomes very important. Substances that can displace these adhering materials are called *wetting agents*. Wetting agents lower the interfacial tensions by preferential adsorption and permit thereby the wetting of a surface by a liquid. When the wetting involves also a dispersion of the displaced film, we have a *detergent*, i.e., a substance suitable for removal of dirt and grime. Typical among the wetting agents used commercially are aliphatic alcohols, sulfonated higher alcohols, sulfonated alkyl naphthalenes, and soaps of various kinds.

COLLOIDS

Introduction. Attention was directed at the beginning of the chapter to the fact that under certain experimental conditions it is possible to distribute one substance through another in particles considerably larger than molecular or ionic size. A clear distinction between true solutions and such dispersions was first made by Thomas Graham in 1861. During a study of diffusion in solutions of various solutes he found that some, such as those present in solutions of salts, acids, bases, and the simpler organic compounds, diffused through a parchment membrane relatively rapidly, while others, such as solutions of gelatin, glue, and silicic acid, diffused extremely slowly. To differentiate these two classes of substances he called the first *crystalloids* or true solutions, and the second *colloids*, i.e., solutions similar to glue.

Graham supposed, as did others after him, that colloids constitute, next to gases, liquids, and solids, a fourth state of matter. However, it is known now that colloids are not a distinct state of matter, but rather a state of *subdivision* in which the particle size is the determining factor as to whether a substance is in true solution, colloidal dispersion, or in coarse suspension in a medium. To substantiate this point of view, stable colloidal dispersions have been prepared of substances that are normally recognized as being crystalline, such as silver chloride, barium, sulfate, gold, and many others. Further, x-ray investigation of various colloidal systems has revealed that frequently the colloidal particles themselves are crystalline and no different in character, except for particle size, from other substances of this type. For these reasons a sharp demarcation between coarse dispersion and colloids on the one hand, and colloids and true solutions on the other, is difficult, and an arbitrary division is made along the following lines. All dispersions in which the particle diameter is less than 1 mμ,[1] i.e., 10^{-7} cm, are taken as true solutions, and these behave as described in the preceding chapters. On the other hand, dispersions containing particles whose diameter exceeds

---

[1] 1 μ (micron) $= 10^{-6}$ meter $= 10^{-4}$ cm.
  1 mμ (millimicron) $= 10^{-3}$ micron $= 10^{-7}$ cm $= 10$ Å (angstroms).

about 0.2 μ, or 0.2 $\times$ 10^{-4} cm, are considered to be merely suspensions
of coarse particles. These are visible under an ordinary microscope,
tend to settle out readily, and are retained by common filter papers.
Finally, dispersions with particles falling between these extremes, i.e.,
between 1 mμ and 0.2 μ, are classified as colloids. These particles are
beyond the range of the ordinary microscope, are not retained by ordi-
nary filters, and in general differ from the other two groups in a number
of properties. An interesting consequence of this classification is that
dispersions of single molecules of high molecular weight substances, such
as proteins, starches, celluloses, and resins, frequently fall within the
colloid range and are, therefore, considered to be colloidal.

Unlike true solutions, which are homogeneous, colloidal systems are
considered to be heterogeneous, with the medium as one phase and the
dispersed substance the other. The medium in which the dispersion
takes place is called the *disperse medium*, the substance dispersed is
called the *disperse phase*, while the complete colloidal solution is referred
to as a *disperse system*. These terms are analogous to solvent, solute,
and solution in true solutions.

Types of Colloidal Dispersions. Since both the disperse phase and
disperse medium may be solid, liquid, or gaseous, colloids may be clas-
sified, like solutions, according to the state of the two to yield the nine
types of disperse systems given in Table 4. Actually these reduce to
eight, since no colloidal dispersion of one gas in another has ever been
observed. The names generally ascribed to some of these types as well
as examples of each category are included.

TABLE 4

Types of Colloidal Dispersions

Disperse Phase	Disperse Medium	Name	Examples
Solid	Gas	Aerosol	Smokes
Solid	Liquid	Sol or Suspensoid	AgCl, Au, As$_2$S$_3$, or S in H$_2$O
Solid	Solid	—	Glasses colored with dispersed metals (ruby glass)
Liquid	Gas	Aerosol	Fogs, mists, clouds
Liquid	Liquid	Emulsion or Emulsoid	Dispersions of H$_2$O in oil or oils in H$_2$O
Liquid	Solid	Gels	Jellies, minerals with liquid inclusions (opal)
Gas	Gas	—	Unknown
Gas	Liquid	Foam	Whipped cream
Gas	Solid	—	Pumice stone

Of the eight types of colloids observed, those of particular importance and interest are the sols or suspensoids, emulsions, gels, aerosols, and foams. However, as an extended discussion of all these would take us too far afield, further reference to the last two will be omitted. Any student interested in these should consult a book such as that of Thomas listed at the end of the chapter. The remaining three classes, namely, sols, emulsions, and gels, will be described here in some detail.

Sols and Their Preparation. Dispersions of solids in liquids can be subdivided broadly into *lyophobic* sols and *lyophilic* sols. Lyophobic (solvent hating) sols are those colloidal dispersions in which there is very little attraction between disperse phase and medium, as in dispersions of various metals and salts in water. Lyophilic (solvent loving or attracting) sols, on the other hand, are dispersions in which the disperse phase exhibits a definite affinity for the medium, and as a result extensive solvation of the colloidal particles takes place. Examples of lyophilic sols are dispersions of gelatin, glue, and other proteins in water. For dispersions in water the terms *hydrophobic* and *hydrophilic* are generally substituted for the more inclusive designations.

As colloids represent a state of subdivision of matter intermediate between true solutions and coarse particles, they may be prepared either by *condensation* of particles in true solution or by *dispersion* of coarse particles into smaller aggregates. Within each of these general methods of preparation are a number of specific means of bringing about colloid formation, and these are listed below:

(a) Condensation Methods	(b) Dispersion Methods
1. Double exchange	1. Mechanical disintegration
2. Reduction	2. Peptization by ions
3. Oxidation	3. Peptization by washing
4. Hydrolysis	
5. Change of solvent	
6. Electric arc	

The *double exchange* method involves essentially the formation of a substance by the exchange of groups, such as the double decomposition of silver nitrate with a chloride, barium nitrate with a sulfate, arsenious chloride with hydrogen sulfide, or a silicate with an acid to form silicic acid. Under ordinary precipitation conditions, i.e., in fairly concentrated hot solutions and in presence of excess reagent, these substances tend to form precipitates. However, in dilute solutions, and usually in the cold and in absence of large excess of electrolyte, these substances can be obtained in a colloidal form. Thus, whereas passage of hydrogen sulfide into a hot acid solution of an arsenious salt yields a precipitate of arsenious sulfide, the same reaction in the cold and in absence of free acid produces a yellowish colloidal sol of the sulfide in water.

Colloidal dispersions of many metals can be obtained by *reduction* of salts of these metals in water with nonelectrolyte reducing agents such as hydrogen, carbon monoxide, phosphorus, formaldehyde, and various hydrazines. Frequently the presence of a trace of electrolyte or a substance like gum arabic or glue is required to stabilize the colloid. By these means sols of silver, gold, platinum, iridium, tellurium, and other metals have been prepared.

Sols of sulfur and selenium have been prepared by *oxidation* of solutions of compounds of these substances, particularly the hydrides H_2S and H_2Se, with air, oxygen, or sulfur dioxide. The student may recall from his experience in qualitative analysis that solutions of hydrogen sulfide on prolonged exposure to air tend to become turbid, although no precipitate is observed. This turbidity is due to colloidal sulfur formed by oxidation of the sulfide by the oxygen in the air.

Colloidal dispersions of many metal oxides can be obtained by *hydrolysis* of solutions of their salts. Thus, a dilute solution of ferric chloride, originally pale yellow, turns a deep red on prolonged boiling to yield a sol of hydrated ferric oxide. Similarly this colloid can be prepared by adding drop by drop concentrated ferric chloride to boiling water. By this means sols of the oxides of chromium, aluminum, cerium, thorium, bismuth, and tin have been obtained. All these metals yield weak bases, and hence hydrolysis of solutions of their salts can be accomplished readily.

When a substance is soluble in one solvent and insoluble in another, addition of the second solvent to a solution of the substance in the first, or vice versa, will frequently result in a colloidal dispersion of the solute. Thus sols of sulfur or phosphorus are formed when solutions of these in alcohol are poured into an excess of water. This *exchange of solvent* is used commercially to make "canned heat" by adding alcohol to a saturated aqueous solution of calcium acetate. The resulting sol eventually absorbs all of the medium and sets to a fairly rigid jelly, the form in which the product is marketed.

Finally, G. Bredig in 1898 showed that colloidal dispersions of metals in water can also be obtained by use of a *direct current electric arc*. When two wires of a metal in question are immersed in water containing a trace of electrolyte and an arc is struck between them, the metal leaves the wire in extremely fine particles and is distributed through the medium as a colloid. It is believed that the metal actually vaporizes in the arc and the vapor is then condensed into colloidal particles. For this reason the electric arc method is classified as a condensation method rather than one of dispersion, which at first glance it may appear to be. Using this method Svedberg succeeded in preparing sols of metals and nonmetals not only in water, but also in a number of organic liquids.

The process of reducing the size of coarse particles to colloidal dimen-

sions with subsequent dispersion is frequently referred to as *peptization*. Among the dispersion or peptization methods the one of *mechanical disintegration* is of considerable importance. In this method the coarse particles are reduced in size by mechanical grinding or shearing until they are sufficiently small to be dispersed in a medium. Usually the substance to be dispersed is mixed with the liquid along with some agent to stabilize the colloid, and the mixture is passed repeatedly through a suitable grinder until part of the solid remains in more or less permanent suspension. For this purpose the *colloid mill* is particularly effective. One form of this device consists essentially of two closely fitted concentric cones which revolve at high speed in opposite directions. The material to be dispersed, along with the medium and protective agent, is fed between the cones where it is sheared apart and dispersed. The colloid mill is also very convenient for preparing emulsions of various types.

Peptization can be accomplished as well in some cases by dilute *solutions of electrolytes*, particularly if they have an ion in common with the solid. Thus precipitates of hydrated ferric oxide can be peptized by solutions of ferric chloride, arsenious sulfide by solutions of hydrogen sulfide, silver chloride by silver nitrate or sodium chloride, etc. Again, peptization can also be brought about frequently by thorough *washing* to remove any adsorbed ions on a precipitate. Freshly precipitated silver chloride or hydrated alumina, if washed long enough with distilled water, will peptize to form sols of these substances in water.

Purification of Colloids. In preparing suspensoids by some of these methods sols result which contain besides the colloidal particles appreciable quantities of electrolyte. To obtain the pure colloid this electrolyte has to be removed. For this purification of a sol three methods are available, namely, (a) dialysis, (b) electrodialysis, and (c) ultrafiltration. In *dialysis* the sol is purified by permitting the electrolyte to diffuse through a porous membrane, such as parchment, cellophane, or collodion, which is permeable to molecules of solvent and substances in true solution, but not to colloidal particles. The usual procedure for dialyzing a sol is to enclose the sol in a sack made of the material chosen for a dialyzing membrane, and then to immerse the membrane in distilled water. On prolonged contact with the water the ions in solution pass through the membrane into the water; but the colloidal particles, not able to pass through the membrane, remain behind to yield a pure sol. Since the rate of diffusion of the electrolyte depends on the difference in concentration of the electrolyte in the sol and in the water surrounding the membrane, it is necessary to avoid an accumulation of electrolyte in the latter by frequent change of the water. In fact, sometimes a continuous stream of water is kept passing about the membrane, assuring a maximum diffusion gradient.

Dialysis is, at best, a very slow process, requiring days, and occasionally weeks, for its completion. It may be accelerated by heating the water and the sol, provided such heating does not lead to precipitation of the colloid. Dialysis is usually not carried to a point where all of the electrolyte is removed, since experience has shown that highly dialyzed sols, by losing too much electrolyte, are not very stable and tend to precipitate readily.

In *electrodialysis* the dialyzing process is accelerated by applying a potential difference across the membrane. Under the influence of this applied field ions migrate faster, and are thereby removed from a sol much more readily. A simple arrangement for conducting an electrodialysis is shown in Fig. 5. The sol is enclosed in a glass tube fitted with membranes M and M' at the tube ends. This tube is immersed in a trough or beaker containing distilled water, and a pair of electrodes A and A' are inserted close to the membranes. On connecting the electrodes to a battery, a current passes through the sol, carrying positive ions through M to electrode A, and negative ions through M' to A'. A stream of water circulating through the outer container removes these ions from the system.
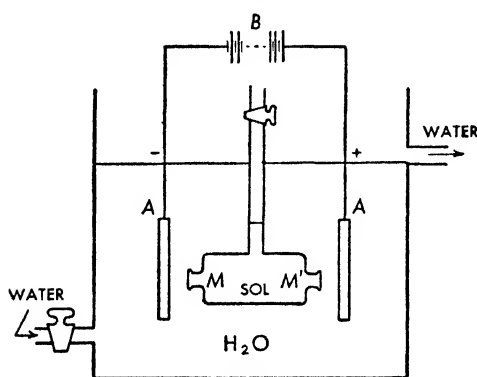
Fig. 5. Electrodialysis

Ultrafiltration is a process similar to filtration of an ordinary precipitate, except that a membrane is used which will permit passage of electrolytes and medium but not of colloid. For this purpose ordinary filter papers are too coarse. The diaphragms mentioned previously may be employed, or a filter may be prepared by impregnating ordinary filter paper with collodion. Other filter media which have been employed are unglazed porcelain and finely sintered glass. Since as a rule ultrafiltration proceeds very slowly, pressure or suction is used to speed the process. By these means it is possible to separate, in the form of a slime, colloid particles from media containing electrolytes, and to remove bacteria from suspensions. These slimes may be suspended, then, in pure media.

Properties of Suspensoids. The properties of suspensoids or sols are best considered under a number of specific heads, namely: (*a*) physical, (*b*) colligative, (*c*) optical, (*d*) kinetic, and (*e*) electrical properties.

The *physical properties* of sols depend on whether the sols are lyophobic or lyophilic. For lyophobic sols such properties as density, sur-

face tension, and viscosity are not very different from those of the medium. This is to be anticipated, since such sols, in being dilute and in showing very slight interaction between suspended matter and medium, affect the properties of the latter little. On the other hand, lyophilic sols show a high degree of solvation of the suspensoid, and as a consequence the physical properties of the medium are modified. This is particularly noticeable in the viscosity, which is much higher for the sol than for the medium. Furthermore, the surface tension of the sol is frequently lower than that of the pure medium.

Sols, like true solutions, exhibit *colligative properties*, but the effects observed are *very much smaller* than for true solutions. In fact, with the exception of osmotic pressure, the effects are practically negligible. The reason for this lies again in the difference in particle size of the two types of dispersions. As was shown in Chapter VI, the magnitude of an observed colligative property depends, at least in dilute solutions, on the number of particles of solute present in a given weight of solvent. When, say, 0.01 mole of solute is dissolved in 1000 grams of solvent to form a true solution, the number of particles resulting is 0.01 N, where N is Avogadro's number. However, if the same amount of substance is colloidally dispersed in the same amount of medium to form aggregates containing 1000 molecules per particle, the number of particles present will be only 1/1000 of that present in the true solution, and hence the colligative effect will be reduced by the same amount. In freezing point lowering, boiling point elevation, and vapor pressure lowering such a decrease in number of particles leads to practically no observable difference between the properties of the sol and the pure medium. But the osmotic pressure, being larger in magnitude than the others, still manifests itself, although again considerably reduced below that of a true solution of the same concentration.

Attempts have been made to estimate the molecular weights of colloidal particles, i.e., the mass of Avogadro's number of particles, by measuring the osmotic pressures of sols and then using the relation $\Pi V = (W/M)RT$ developed in Chapter VI. Though these methods give some idea of the masses of colloidal particles, they are open to considerable question, particularly because it is never certain in such experiments that all electrolyte is absent. Presence of even small amounts of electrolyte may lead to osmotic effects larger than those due to the colloid, and may vitiate thereby any significance to be ascribed to the results.

Optical Properties of Colloids. A true solution is optically empty. A beam of light passed through the solution is partly absorbed and partly transmitted, but there is no tendency of the solution to scatter the light. On the other hand, a beam directed at a colloidal dispersion has its

path through the sol illuminated because the light is scattered by the colloidal particles. This tendency of colloids to scatter light is called the *Tyndall effect*, while the illuminated path is referred to as a *Tyndall beam* or cone. All of us are familiar with the Tyndall beams shown by dust particles in air when a ray of sunlight enters a darkened room, or with the beam thrown by searchlights or headlights of a car on a foggy night. These are all due to scattering of light either by dust particles or by water droplets in colloidal suspension in air.

Scattering of light is observed only when light strikes a medium which is heterogeneous with finely divided particles. Further, Lord Rayleigh showed that the intensity of the scattered light depends on the difference in the refractive indices of the disperse medium and the disperse phase, and is greater the larger the difference. For this reason the Tyndall effect is particularly well defined in lyophobic colloids, where the difference in refractive indices of the two phases is pronounced. However, in lyophilic colloids, where extensive solvation of the disperse phase tends to obliterate to an appreciable extent the difference in refractive indices, the Tyndall effect may be weak and difficult of observation.
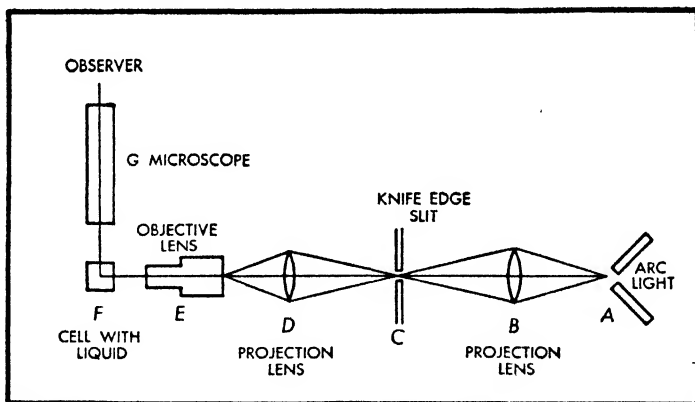
Fig. 6. Schematic Arrangement of Slit Ultramicroscope

The ability of colloids to scatter light is the basis of the instrument invented by Siedentopf and Zsigmondy in 1903 for making colloidal particles visible. A schematic diagram of this instrument, called an *ultramicroscope*, is shown in Fig. 6. A beam of light *A*, from an arc lamp or some other strong source, is condensed by the projection lens *B* onto a knife-edge slit *C*, whose function it is to control the width of the beam. This beam is focused by means of projective lens *D* onto objective lens *E*, which reduces the image of the slit appreciably before permitting the light to enter the cell *F* containing the colloid. The light scattered by the colloidal particles is observed then through a microscope *G*, located

at an angle of 90° to the original path of the beam. With this arrangement the colloidal particles appear as pin points of light moving about against the dark background of the nonscattering disperse medium. An ultramicroscope such as the one described can resolve particles as small as 5 $m\mu$, and with certain refinements a resolution as low as 3 $m\mu$ has been claimed.

The ultramicroscope permits not only the observation of colloidal particles, but also an estimation of their size. First, it should be emphasized that the size of the pin point of light seen in the ultramicroscope is not an indication of the actual diameter of the particles. However, it is possible to deduce the dimensions of colloidal particles as follows. The mass of disperse phase present in a given volume of sol is first determined by some suitable means, such as evaporation of the sol and weighing of the residue. Next the sol is placed in the ultramicroscope, and the number of particles present in a definite volume, as determined by means of an ocular micrometer, is counted. Assuming now that all particles have the same size, that they are spherical in shape, and that their density is the same as for the disperse phase in bulk, the volume of a single particle V follows as

$$V = \frac{m}{nd} \tag{10}$$

where m is the mass of particles present in the volume in which n particles were found, while d is their density. Remembering that the volume of a sphere is $4/3\,\pi r^3$, the average diameter of the particles l follows from equation (10) as

$$l = 2\sqrt[3]{\frac{3}{4}\frac{V}{\pi}}$$
$$= 2\sqrt[3]{\frac{3\,m}{4\,\pi nd}} \tag{11}$$

Using this method Burton[1] showed that in dispersions of noble metals in water in his experiments the particle size ranged from 0.2 to 0.6 μ in diameter.

Kinetic Properties of Sols. There are several properties of sols which can be discussed conveniently under this heading, since they all involve an application of kinetic theory to colloidal dispersions. These are diffusion, Brownian motion, and sedimentation.

Attention has already been directed to Graham's observation that colloidal particles diffuse much more slowly than solutes in true solution. Albert Einstein[2] and others showed that it is possible to derive an ex-

[1] Burton, Phil. Mag., **11**, 425 (1906).
[2] A. Einstein, Ann. Physik, **17**, 549 (1908); **19**, 371 (1906); Z. Elektrochem., **14**, 235 (1908).

pression for the diffusion of colloidal particles in a medium provided it is assumed that the kinetic theory as expressed in the van't Hoff equation for osmotic pressure applies, and that the colloid particles are spherical and large compared to the molecules of the medium. With these assumptions the equation for the diffusion coefficient D, i.e., the amount of colloid diffusing across unit area per unit time under a concentration gradient of unity, follows as [1]

$$D = \frac{RT}{N}\left(\frac{1}{6\,\pi\eta r}\right) \tag{12}$$

where R is the gas constant, T the absolute temperature, N Avogadro's number, η the viscosity of the medium, and r the radius of the colloidal particles. This equation has been employed to estimate from diffusion measurements either r, when N was assumed, or N, when r was known, and has been found to hold well for various types of dispersions. However, the usual practice of testing this relation is in conjunction with Brownian movement, and this is to be discussed now.

Colloidal particles observed in the ultramicroscope are found to be not stationary, but in a ceaseless random and swarming motion. This kinetic activity of particles suspended in a liquid, first observed by the botanist Brown in 1827, is named after him *Brownian motion*. Careful investigation of the effect has definitely established that Brownian motion is not due to external influences such as vibration, heat, light, gravitational, electric, or magnetic forces, but is intimately associated with molecular motion of the molecules of the medium. Because of bombardment of the dispersed particles by molecules of the medium, the particles acquire at equilibrium a kinetic energy equal to that of the medium at the given temperature. Since the particles in dispersion are considerably heavier than the molecules of the medium, this acquired kinetic energy manifests itself in a Brownian movement which is considerably slower than that of the molecules of the medium, and may thus be observed in the ultramicroscope.

The validity of this explanation of Brownian motion is borne out by mathematical considerations again due to Einstein and found in the references cited. Einstein showed that, on the basis of the assumptions mentioned in connection with diffusion, the diffusion coefficient of a colloid should be related to the *average displacement* Δ, produced by Brownian movement in time t along the x axis by the equation

$$D = \frac{\Delta^2}{2\,t} \tag{13}$$

[1] For derivation see either references cited, or various books on colloids, such as the one by Thomas (pp. 49–51).

Eliminating D between equations (12) and (13), the expression for the square of the average displacement along the x axis follows as

$$\Delta^2 = \frac{RT}{N}\left(\frac{t}{3\,\pi\eta r}\right) \tag{14}$$

This equation was used by Perrin in his classical experiments on suspensions of gamboge and mastic in water for the determination of Avogadro's number, N. The displacements Δ were estimated by charting with the aid of the ultramicroscope the positions of a specific particle at various time intervals and calculating therefrom the average horizontal displacements. The radius of the particles was determined from sedimentation experiments in a manner to be described shortly. These data were substituted then in equation (14) to calculate N. From a large number of experiments with particles ranging from 0.2 to 11.5 μ Perrin found N to be 6.85×10^{23}, as compared with the present best value of 6.02×10^{23}. Considering the difficulties involved in these experiments, Perrin's work constitutes an excellent confirmation both of the applicability of the kinetic theory to colloids and of the postulated nature of Brownian motion as expressed in equation (14).

Sedimentation of Suspensoids. Although colloidal dispersions may be stable over long periods of time, sometimes years, they nevertheless tend to settle out slowly under the influence of gravity on prolonged standing. The rate of settling can usually be followed quite readily, for the boundary between clear medium and sol is, as a rule, quite distinct. Since from the observation of such rates it is possible to arrive at the dimensions of the suspended particles and their masses, the principles of the process will be developed here in some detail.

For this purpose consider a tube such as that shown in Fig. 7 containing a disperse phase of density d suspended
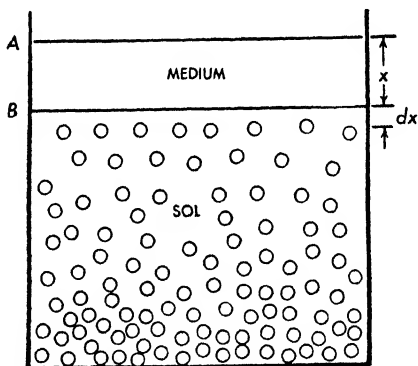
Fig. 7. Sedimentation Under Gravity

in a medium of density d_m and viscosity η. If d is greater than d_m, the particle will settle; otherwise, the particle will be displaced upward. Any particle of radius r tending to settle under gravity will be opposed by the frictional force of the medium. Now, it was shown in Chapter III, equation (51), that when equilibrium is established between these two forces a spherical particle will fall according to Stokes's law with

a constant velocity v, given by

$$v = \frac{2\,r^2 g(d - d_m)}{9\,\eta} \tag{15}$$

where g is the acceleration of gravity. Since v at a point such as x from the top, A, is obviously dx/dt, the rate at which the height of level B changes with time, equation (15) may be written as

$$\frac{dx}{dt} = \frac{2\,r^2 g(d - d_m)}{9\,\eta} \tag{16}$$

Integrating this equation between the limits $x = x_1$ at $t = t_1$, and $x = x_2$ at $t = t_2$, we obtain

$$(x_2 - x_1) = \frac{2\,r^2 g(d - d_m)(t_2 - t_1)}{9\,\eta} \tag{17}$$

By knowing d, d_m, and η, and by measuring the distances x_2 and x_1 at two different times t_2 and t_1, it is possible to calculate from equation (17) the radius of the settling particles. This is the procedure Perrin used in his experiments. Once the radius is known, the mass of a single particle follows from $m = 4/3\,\pi r^3 d$, and the molar weight of the disperse phase follows from $M = Nm$.

Unless the suspended particles are large, sedimentation under the influence of gravity is an extremely slow process. However, it is possible to accelerate greatly the sedimentation by means of *ultracentrifuges* developed by The Svedberg and his coworkers. An ultracentrifuge is essentially a very high-speed centrifuge in which the centrifugal force of rotation is substituted for the force of gravity. By whirling colloidal dispersions in cells placed in specially designed rotors, forces as high as one million times that of gravity have been achieved. Under these conditions even finely dispersed sols can be sedimented in a relatively short time. To observe the rate of settling, optical methods have been developed which permit photographing the level of the sol at various stages of settling without stopping the ultracentrifuge.

Since the force operating on any particle at a distance x from the axis of rotation of the ultracentrifuge is $\omega^2 x$, where ω is the angular velocity of rotation, substitution of this quantity into equation (16) in place of g gives immediately the rate of settling, namely,

$$\frac{dx}{dt} = \frac{2\,r^2 \omega^2 x(d - d_m)}{9\,\eta} \tag{18}$$

Separating the variables, and integrating between the same limits as before, we have

$$\ln \frac{x_2}{x_1} = \frac{2\,r^2 \omega^2(d - d_m)}{9\,\eta}\,(t_2 - t_1) \tag{19}$$

Through equations (19) r may again be calculated from the values of x and t at two stages of the sedimentation and from a knowledge of the speed of rotation of the ultracentrifuge, ω. This procedure for obtaining r, and therefrom the molar weight of a suspensoid, is called the *sedimentation velocity* method.

An alternate procedure for obtaining r or the molar weight of a disperse phase is the *sedimentation equilibrium* method. If a suspensoid is whirled sufficiently long in an ultracentrifuge, a stage is reached at which the sol no longer settles. At this point an equilibrium is reached between the rate of sedimentation and the rate at which the suspended particles tend to diffuse back into the more dilute portions of a cell against the centrifugal force. By equating these two rates it is possible to arrive at the distribution of concentration with distance in various parts of the sedimentation cell. This distribution is given by the relation

$$\ln \frac{C_2}{C_1} = \frac{M\omega^2(d - d_m)(x_2^2 - x_1^2)}{2\,RTd} \tag{20}$$

where C_1 and C_2 are the concentrations of colloid at levels x_1 and x_2, while $M = (4/3\,\pi r^3)Nd$ is the weight of Avogadro's number of particles, or *the molecular weight of the colloid*. By determining the concentrations C_1 and C_2 at the two levels x_1 and x_2 in the settling cell at sedimentation equilibrium, M can readily be calculated from equation (20).

TABLE 5

MOLECULAR WEIGHTS OF PROTEINS AS DETERMINED BY
USE OF ULTRACENTRIFUGE [1]

Protein	Source	Molecular Weight	r in $m\mu$
Egg albumin	Hens' eggs	34,500	2.17
Insulin	Beef pancreas	35,100	2.18
Hemoglobin	Horse blood	68,000	Not spherical
Serum albumin	Horse blood	67,500	Not spherical
Serum globulin	Horse blood	103,800	Not spherical
Amandin	Almonds	208,000	3.94
Cocosin	Coconut	208,000	3.95
Edestin	Hemp seed	212,000	3.94
Hemocyanin	Blood of octopus	2,000,000	Not spherical
Hemocyanin	Blood of Helix	5,000,000	12.0

Extensive experiments with the ultracentrifuge have been carried out on various dispersions, particularly proteins, by Svedberg and others. Table 5 gives the molecular weights of some proteins in water dispersion

[1] From Thomas, "Colloid Chemistry," McGraw-Hill Book Company, Inc., New York, p. 353.

obtained in this manner, as well as the radii of some of the molecules in $m\mu$. From this table two things are apparent, namely, (1) that the molecular weights of colloidal substances may be extremely large, and (2) that the molecular weights of the higher proteins appear to be multiples of a unit equal to that of the simplest, or 35,000. The latter fact is of importance in theories of protein structure.

Electrical Properties of Sols. Colloidal dispersions possess a number of electrical properties which are intimately associated with their ability to adsorb from solution ions, molecules of medium, or both. Solid in bulk has a relatively small surface area for a given mass. Thus the area of a cube 1 cm on edge is only 6 sq cm. If this cube is now subdivided into cubes 10 $m\mu$ per edge, the order of colloidal dimensions, 1 cc of matter will yield 10^{18} cubes with a total surface area of 6×10^6 sq cm. With such a tremendous development of surface adsorption phenomena should become very pronounced, and this is actually the case with colloidally dispersed particles. Evidence accumulated from the study of colloids has established that lyophobic sols adsorb primarily ions of electrolyte from the solutions in which they are prepared. On the other hand, the particles in lyophilic sols attract onto themselves first a layer of medium to which, depending on the conditions existing in the solution, may or may not be adsorbed ions. Although the nature of the dispersoid particles determines whether a colloid is to be lyophobe or lyophile, it is the character of the adsorbed phase on the final particle, called a *micelle*, which eventually controls the stability of the sol and establishes the manner in which the particles are to behave in an electric field.

Any solid in contact with a liquid tends to develop a difference in potential across the interface between the two. Thus it has been shown that, when water is brought into contact with a glass tube, the tube takes on a negative potential with respect to the water. The explanation advanced for this observation is that the glass adsorbs hydroxyl ions from the water and becomes negatively charged. To counterbalance this charge, hydrogen ions are attracted to the surface to form a double layer of charges, with the negative charges on the glass and positive charges in the water immediately adjacent to these, as shown in Fig. 8(*a*). Such an arrangement of charges, called a *Helmholtz double layer*, leads to a difference of electric potential between the solid and liquid.

In the original concept of the double layer the charges next to the surface were considered to be fixed, while the compensating charges in the liquid, along with the latter, were thought to be movable. However, more recent considerations point to a somewhat modified picture of the double layer. Instead of postulating that all the balancing charges are held in a single movable layer near the surface, it is believed, rather, that the layer is more diffuse in character and extends partway into the

liquid in the manner illustrated in Fig. 8(b). Here the charges shown immediately adjacent to the surface are probably on the surface, with some of the compensating charges held in a stationary liquid layer adhering to the surface. The remaining charges are distributed then through the electrolyte next to this layer in the form of a diffuse and movable atmosphere analogous to the one postulated by Debye and Hückel for electrolytes. In terms of this concept the potential drop between the surface and the liquid does not occur any more across the fixed double layer, but is involved in two parts, (1) between the surface and the stationary compensating layer, and (2) between the latter and the body of the solution. The sec-

HELMHOLTZ DOUBLE LAYER

(a)

DIFFUSE DOUBLE LAYER

(b)

Fig. 8. Helmholtz and Diffuse Double Layers

ond of these potential drops, called the *electrokinetic* or *zeta potential*, is the one involved in various nonstatic electrical properties of solid-liquid interfaces, and is the one responsible for the electrical effects observed in colloids.

Since colloidal dispersions of solids in liquids produce solid-liquid interfaces, it may be anticipated that as a result of differential adsorption of ions by the solid a diffuse double layer of the type described above will be established at the particle boundary, and this will lead to the appearance of a zeta potential. Further, such adsorption must result in an electric charge on the particle of sign equal and opposite to that of the surrounding medium; i.e., both the particles and medium become electrically, though oppositely, charged. If this be the case, both the particle and the medium should respond to applied electric fields, and should migrate under these in opposite directions. This is actually so. When the experimental conditions are adjusted so as to permit the migration of dispersoid but not of medium, we encounter the phenomenon of *electrophoresis*. On the other hand, when migration of dispersoid, but not that of medium, is prevented, we have *electroosmosis*.

Electrophoresis. The migration of electrically charged colloidal particles under an applied electric potential is called *electrophoresis*. This effect can be followed readily in the apparatus shown in Fig. 9, called a *Burton tube*. This apparatus consists of a U-tube fitted with a stopcock for drainage, and a funnel-shaped filling tube with stopcock (not shown) attached to the back of the U-tube. A solution of some suitable electro-

lyte of lower density than the sol is first placed in the tube, and the sol next introduced through the funnel so as to displace the electrolyte upward and produce sharp boundaries in the two arms. Electrodes are then inserted, as shown, and connected to a source of potential, such as a high-voltage battery. When the colloidal particles are negatively charged it is observed that the level of the sol falls gradually on the negative electrode side and rises simultaneously on the positive side, i.e., the colloidal particles move toward the positive electrode. Conversely, when the particles are positively charged, the reverse occurs. If the electrophoresis is permitted to proceed until the appropriate electrode is reached, the sol will discharge and precipitate.

Fig. 9. Burton Tube for Electrophoresis

By noting whether colloid particles migrate in an electric field and the direction of migration, it is readily possible to determine whether the particles in a given colloid are charged, and what is the sign of the charge. Through such means it has been established that lyophobic colloids bear charges that are of the same sign for a given colloid, i.e., lyophobic colloids are specific in their adsorption of either positive or negative ions. Thus sulfur, metallic sulfide, and noble metal sols are found to migrate to the positive electrode, and are, therefore, negatively charged. On the other hand, metal oxide sols, such as iron and aluminum oxide, are positively charged. This same specificity of charge is observed also with some lyophilic sols. However, with certain sols of the latter type, particularly the proteins, the sign of the charge depends on the pH (hydrogen ion concentration) of the solution. Above a certain pH value, characteristic for each sol, the particles are negatively charged, while below this pH they have a positive charge. At the pH value in question, called the *isoelectric point*, the particles are uncharged, and consequently do not migrate in an electric field. As a rule the isoelectric point does not come at a definite pH but covers a pH range, being, for instance, 4.1 to 4.7 for casein from human milk, and 4.3 to 5.3 for hemoglobins from various sources.

Electrophoresis can also be utilized for quantitative measurement of the actual rate with which colloid particles migrate. By determining the time necessary for a sol to migrate a definite distance under a potential difference applied over a given length of conducting path, the rate in

centimeters per second under a potential drop of 1 volt per centimeter can be calculated. Such calculations show that the migration velocities of colloidal particles are not much different from those of ions under the same conditions, and are of the order of 10 to 60 $\times$ 10^{-5} cm per second per volt per centimeter.

Electroosmosis. When electrophoresis of a dispersoid in an applied electric field is prevented by some suitable means, the medium can be made to move under the influence of the applied potential. This phenomenon is referred to as *electroosmosis*, and is a direct consequence of
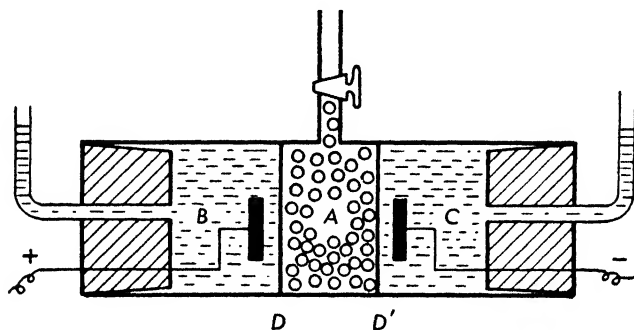
Fig. 10. Electroosmosis

the existence of a zeta potential between the colloid particles and the medium. Figure 10 shows a simple apparatus for observing electroosmosis. The colloid is placed in compartment A, separated from compartments B and C by dialyzing membranes D and D'. B and C are filled with water up to the marks indicated on the sidearms. When a potential is applied across the two electrodes placed close to the membranes in B and C, the liquid level is observed to fall on one side and rise on the other owing to passage of water through D and D'. The direction of flow of the water depends on the charge of the colloid. For positively charged sols, the medium is negatively charged, and hence the flow will take place from C to B. For negatively charged colloids the reverse will be true, and the level on the C side will rise.

Stability of Colloids. There is one question of considerable interest and importance in connection with colloids which we have not as yet considered, namely: Why are colloidal dispersions stable? The answer can be found for lyophobic colloids in their electric charge, for lyophilic colloids in both their electric charge and their extensive solvation. To bring about precipitation of a dispersion the particles must coalesce on collision into aggregates large enough to settle out. In lyophobic sols such agglomeration is prevented by the electric charges on the particles. As we have seen above, all particles in a lyophobic sol bear charges of

the same sign, and hence when such particles approach each other they are repelled by the electrostatic forces between them. The result is that as long as these charges are present each particle is protected from every other by an electric field, and no approach close enough for coalescence into larger aggregates is possible. In lyophilic colloids the same type of electrostatic repulsion is operative, along with the added factor of solvation. The colloidal particle becomes encased in a sheath of solvent which acts as a further barrier to contact between aggregates of disperse phase. Removal of the surface charge in lyophobic colloids would be sufficient to bring about coagulation. In lyophilic colloids however, this may decrease stability, but it does not necessarily result in precipitation. This is well illustrated by proteins at their isoelectric point, where the micelles are uncharged. Although in this state the protein dispersions are least stable, they do not always precipitate because they are still protected by a surrounding layer of water. However, as soon as this layer is removed by some appropriate means, coagulation takes place, and the protein settles as a precipitate.

The added stability of lyophilic colloids resulting from their ready solvation is frequently utilized to increase the stability of lyophobic sols. As will be shown, lyophobic colloids may be precipitated by the addition of electrolyte. However, when some lyophilic colloid is added to the lyophobic sol, it is found that the amount of electrolyte necessary for precipitation of the sol is greater, i.e., the stability of the sol has increased. This increase in stability is due to a coating over of the lyophobic particles by lyophilic colloid. As a result the sol acquires the properties of a lyophilic substance, and exhibits the greater resistance of such particles to precipitation. A lyophilic colloid used in this manner to increase the stability of a lyophobic one is called a *protective colloid.*

Precipitation of Sols. To bring about precipitation of a disperse phase, conditions must be created which are opposite to those which lead to stability. For lyophobic sols this means that the charge on the particles must be removed; or, what amounts to the same thing, the zeta potential of a micelle must be reduced to a point where coagulation of dispersed particles and precipitation become possible. On the other hand, to effect precipitation of lyophilic sols, removal of both the charge and the adhering layer of solvent is necessary.

Neutralization of electric charges on lyophobic particles may be accomplished by the use of an applied electric field, such as is used in electrophoresis. The charged particle on coming in contact with an electrode of sign opposite to itself is discharged and precipitated. A more common method of producing precipitation in sols is by the addition of electrolytes, which, though essential in small concentrations for stability, produce in higher concentrations flocculation of the disperse phase. The

TABLE 6

PRECIPITATION OF HYDROSOLS BY ELECTROLYTES

Ferric Hydroxide Sol (positive)		Arsenious Sulfide Sol (negative)	
Electrolyte	Concentration (millimoles/liter)	Electrolyte	Concentration (millimoles/liter)
NaCl	9.3	NaCl	51
KCl	9.0	KCl	50
KI	16.2	HNO$_3$	50
BaCl$_2$	9.6	HCl	31
		MgCl$_2$	0.72
K$_2$SO$_4$	0.20	CaCl$_2$	0.65
MgSO$_4$	0.22	BaCl$_2$	0.69
K$_2$C$_2$O$_4$	0.24	ZnCl$_2$	0.69
K$_2$Cr$_2$O$_7$	0.19	Al(NO$_3$)$_3$	0.095
		AlCl$_3$	0.093
K$_3$Fe(CN)$_6$	0.096	Ce(NO$_3$)$_3$	0.080

amount of electrolyte required to precipitate a given sol depends on the nature of both the sol and the electrolyte added. Schulze first pointed out a rule, confirmed later by others, that for a given sol the precipitating power of an electrolyte is determined primarily by the valence of the ions opposite in sign to that of the colloid. The validity of this rule may be judged from Table 6, where are given some results obtained by Freundlich for the precipitation of a positive and negative sol by various electrolytes. The concentrations refer to the minimum amounts of electrolyte required to produce precipitation in 2 hours. From this table it is readily evident that, although some specificity does exist, the precipitating action of various electrolytes depends essentially on the valence of the anions in the positive sol, and on the valence of the cations in the negative sol. Further, the precipitating power of an electrolyte increases very rapidly with increase in the valence of the anion or cation, the ratios being approximately 1 : 40 : 90 for the ferric hydroxide and 1 : 70 : 500 for the arsenious sulfide.

Frequently a sol may be coagulated by adding to it a colloid of charge opposite to itself. In the process both sols may be partially or completely precipitated. A case in point is the precipitation of the negative arsenious sulfide sol by the positive ferric hydroxide hydrosol. Again, flocculation can occasionally be effected by boiling or freezing. In boiling, the amount of electrolyte adsorbed by the sol is reduced; and, if this reduction is sufficient to produce coagulation, the sol can be destroyed. Freezing, on the other hand, results in removal of medium; and, if carried far enough, there may not be enough medium left to keep the dispersoid suspended.

To produce flocculation in lyophilic sols both the charge and adhering solvent must be removed. This may be accomplished in two ways. First, if solvents such as alcohol or acetone, which have a high affinity for water, are added to a hydrophilic colloid, the disperse phase undergoes a change in character to one approximating more closely that of a lyophobic sol. That is, the dehydration of the micelles by the alcohol or acetone yields particles whose stability is due solely to their charge. A small quantity of electrolyte added to the sol in this condition produces ready flocculation. Again, precipitation can also be brought about by the use of sufficiently high concentrations of certain ions. This "salting out" of lyophilic colloids by high concentrations of electrolyte is due, apparently, to the tendency of the ions to become solvated, causing the removal of adsorbed water from the micelles. The "salting out" power of various ions is given by the *Hofmeister* or *lyotropic series*, in which the ions are arranged in order of *decreasing* precipitating effectiveness. For anions this series gives

$$\text{citrate}^{---} > \text{tartrate}^{--} > SO_4^{--} > PO_4^{---} > \text{acetate}^- > Cl^- >$$
$$NO_3^- > I^- > CNS^-$$
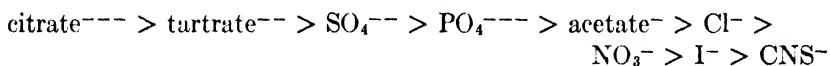
while for cations,

$$Mg^{++} > Ca^{++} > Sr^{++} > Ba^{++} > Li^+ > Na^+ > K^+ > Cs^+$$
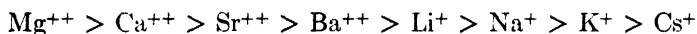
However, this order is not always rigidly valid. Depending on the nature of the colloid being precipitated, the relative positions of various anions and cations may be changed.

Emulsions. A colloidal dispersion of one liquid in another immiscible with it is called an *emulsion*. Emulsions can be prepared by agitating a mixture of the two liquids or, preferably, by passing the mixture through a colloid mill referred to as a homogenizer. Such emulsions prepared from the pure liquids only are generally not stable, and settle out on standing. To prevent this, small quantities of substances called *emulsifying agents* or *emulsifiers* are added during preparation to stabilize the emulsions. These are usually soaps of various kinds, long-chain sulfonic acids and sulfates, lyophilic colloids, or finely divided powders of carbon, metals, or some salts.

If we employ the term *oil* to designate any liquid immiscible with water and capable of forming an emulsion with it, we may classify emulsions into two classes, namely: (1) emulsions of oil in water, in which the disperse phase is the oil while water is the medium, and (2) emulsions of water in oil, in which the functions of the two are reversed. The type of emulsion that results on agitation of two liquids depends on the relative proportions of the two in the mixture. As a rule the one in excess will act as the "outside phase" or medium, the other as the "inner

phase" or dispersoid. The exact excess required for the formation of a particular type of emulsion is hard to define. In fact, Stamm and Kraemer [1] showed that emulsions of oleic acid in water can be prepared as long as the proportion of acid in the mixture does not exceed 40 per cent. Once this percentage is exceeded, an emulsion of water in the acid results. The type of emulsion obtained depends also on the nature of the emulsifying agent used. Water-soluble alkali metal soaps and basic metal sulfates generally favor the formation of emulsions of oil in water. On the other hand, water-insoluble soaps, such as those of zinc, aluminum, iron, and the alkaline earth metals, favor the formation of emulsions of water in oil. Frequently the first of these types of emulsions can be converted to the second by addition of heavy metal ions. Apparently under these conditions a water-soluble emulsifier is converted to one that is water insoluble, and a reversal of phase is brought about.

Whether a particular emulsion belongs to one type or another can be ascertained in several ways. If water is the outside phase, then any water added to the emulsion will be readily miscible with it, while oil will not. Similarly, for oil as the outside phase the miscibility will be with oil, but not water. By observing the behavior on such additions under a microscope, it is readily possible to identify the nature of the emulsion. Another means of distinguishing the two depends on the fact that a small amount of electrolyte added to an emulsion will make the latter conducting if water is the outside phase, but will have very little effect on the conductance if oil is the disperse medium. Therefore, by testing the conductance in presence of the electrolyte the two types of emulsions may be differentiated.

The particle size of the distributed phase is usually larger in emulsions than in sols, ranging from 0.1 to more than 1μ in diameter. For this reason globules of liquid can frequently be seen under a microscope. Otherwise the properties of emulsions are not very different from those of lyophobic sols. They show the Tyndall effect, and Brownian motion can be observed provided the particles are not too large. Again, the globules are generally negatively charged. Like such colloids they migrate in an electric field and exhibit sensitivity to added electrolytes, particularly those containing multivalent positive ions.

Emulsions can be broken to yield the constituent liquids by heating, freezing, severe jarring, centrifuging, by addition of appreciable quantities of electrolyte to salt out the disperse phase, or by chemical destruction of the emulsifying agent. Centrifuging is commonly resorted to for the separation of cream from milk, the separation of water from oil, etc. Because of the negative charge of the globules, electrolytes containing divalent and trivalent positive ions are particularly effective in reducing

[1] Stamm and Kraemer, J. Phys. Chem., **30**, 992 (1926).

the zeta potential and bringing about a salting out of the emulsified substance. The means employed for the destruction of the emulsifying agent depend on its nature. As an example may be mentioned the precipitation of an oil in water emulsion stabilized by a sodium or potassium soap through addition of a strong acid. ' The acid hydrolyzes the soap to liberate the free fatty acid, and, since the latter is not a good emulsifying agent, the emulsion separates. The same result can be accomplished by careful addition of positive ions which tend to bring about a reversal of phase. If an excess is avoided, a segregation of the two phases will take place before reversal of the oil in water to a water in oil emulsion becomes possible.

Gels. Coagulation of a lyophobic or lyophilic sol yields usually a precipitate which may or may not be gelatinous. However, if the conditions are right, it is possible to obtain the disperse phase as a more or less rigid mass enclosing within it all of the liquid. The product in this form is called a *gel*, while the process by which it is formed is called *gelation.*

Depending on their nature, gels may be prepared generally by one of the following three methods: (1) cooling, (2) double decomposition or metathesis, or (3) change of solvents. Gels of agar-agar, gelatin, and other substances of this kind are prepared readily by cooling a not too dilute dispersion of these substances in hot water. As the gel cools, the highly hydrated dispersed particles lose stability, agglomerate into larger masses, and eventually mat together to form a semirigid structure that entraps any free medium. The result is a gel. The second method is illustrated by the formation of silicic acid gels on addition of an acid to a water solution of sodium silicate. The free silicic acid thus liberated is highly gelatinous due to hydration, and sets more or less rapidly, depending on the concentration, to a solid gel. Finally, some gels are formed on changing rather suddenly the solvent in which a substance is dissolved to one in which it is insoluble. An example of this is a gel of calcium acetate. When alcohol is added rapidly to a solution of calcium acetate. in water, the salt is suddenly thrown out of solution as a colloidal dispersion which subsequently sets to a gel containing all of the liquid.

Gels may be subdivided into two kinds, *elastic,* of which agar-agar and gelatin are examples, and *nonelastic,* such as silica gel. The prime difference in these two types of gels is their behavior on dehydration. A completely dehydrated elastic gel can be regenerated by addition of water; however, once a nonelastic gel is freed of moisture, addition of water will not bring about gelation. This difference in hydration behavior seems to lie in a difference in the structure of the dried gels. In elastic gels the fibrils composing the gel are flexible, and on hydration they can expand again to reform the gel. In nonelastic gels, on the other hand, the struc-

ture is much more rigid, and though the gel can take on some water, it cannot expand sufficiently to enclose all of the liquid.

A property of partially dehydrated elastic gels is their ability to imbibe water when immersed in the solvent. The amount of water thus absorbed may be large, leading to an appreciable expansion or *swelling*. Many gels, both elastic and nonelastic, experience also a shrinkage in volume on standing with an attendant exudation of solvent. This process is called *syneresis*. Again, some gels, particularly those of hydrous oxides and gelatin, liquefy readily when shaken to form a sol which on standing reverts to a gel. This sol-gel transformation is reversible, and is referred to generally as *thixotropy*. This same reversibility is observed in gels that can be prepared by cooling, except that the liquefaction is brought about by heating, the gelation by reducing the temperature. The temperature at which the latter type of transformation occurs is usually quite definite and reproducible.

Importance of Colloids. Colloidal dispersions play such an important role in our everyday life and in industry that a further word on their extensive occurrence may be in order. Many of our foods are colloids. Preeminent among these is milk, which is an emulsion of butterfat in water stabilized by casein as an emulsifying agent. Others are salad dressings, clear soups, jello, fruit jellies of various kinds, and whipped cream. Many medicinal preparations are either emulsions or suspensions. Soap, our most useful detergent, is a colloidal electrolyte whose detergent action is ascribed to its ability to emulsify dirt, grime, and surface oils with water, and cause their removal by rinsing.

Latex, from which rubber is made, is a colloidal suspension of negatively charged rubber particles. This fact is taken advantage of to plate rubber by electrophoresis on various metallic objects. Paints, lacquers, varnishes, and enamels are dispersions of paint pigments or gums in suitable vehicles or solvents. Silica and alumina gels find application as sorbents for various gases, and as drying agents. Practically all our new plastics are colloids, as are the celluloses, resins, gums, and glues. Though this list could be continued, it is sufficient to point out that from dispersions of cellulose alone we obtain such important products as rayon, Cellophane, and paper.

Finally, smokes, dusts, clouds, mists, and fogs are all suspensions of either solids or liquids in gases. Although some of these, such as clouds, are of benefit, fogs are a hazard, while dusts and smokes create sanitary and health problems. When practicable or necessary such aerosols are generally destroyed by electrical precipitation in a Cottrell precipitator. This consists of two plates across which a very high potential is applied, and between which the dust, smoke, or mist is passed. The dispersed particles are attracted to the plate of charge opposite to their own, are

discharged, and fall from the plate into containers placed below to receive the precipitate.

REFERENCES FOR FURTHER READING

1. N. K. Adam, "The Physics and Chemistry of Surfaces," second edition, Oxford University Press, New York, 1938.
2. J. Alexander, "Colloid Chemistry," 4 volumes, Reinhold Publishing Corporation, New York, 1926–1932.
3. E. F. Burton, "The Physical Properties of Colloidal Solutions," third edition, Longmans, Green and Company, New York, 1938.
4. H. Freundlich, "Colloid and Capillary Chemistry," Methuen and Co., Ltd., London, 1926.
5. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.
6. R. J. Hartman, "Colloid Chemistry," Houghton Mifflin Company, Boston, 1939.
7. E. A. Hauser, "Colloidal Phenomena," McGraw-Hill Book Company, Inc., New York, 1939.
8. H. N. Holmes, "Introductory Colloid Chemistry," John Wiley & Sons, Inc., New York, 1934.
9. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931, Chapter XX.
10. A. W. Thomas, "Colloid Chemistry," McGraw-Hill Book Company, Inc., New York, 1934.
11. H. B. Weiser, "Inorganic Colloid Chemistry," 3 volumes, John Wiley & Sons, Inc., New York, 1933–1938.

PROBLEMS

1. From their critical temperatures arrange the following gases in the order in which you would expect them to be most readily adsorbed by charcoal at 15° C: Ethane, ethylene, n-pentane, argon, nitrogen, and chlorine.

2. The following data are given for the adsorption of CO on wood charcoal at 0° C. The pressure P is in mm Hg, while x is the volume of gas in cubic centimeters, measured at standard conditions, adsorbed by 2.964 g of charcoal.

P	x	P	x
73	7.5	540	38.1
180	16.5	882	52.3
309	25.1		

By plotting the proper functions, evaluate the constants k and n of the Freundlich equation, and the constants a and b of the Langmuir equation.

3. From the results of the preceding problem calculate the volume of CO at standard conditions adsorbed by 1 g of wood charcoal at 0° C when the partial pressure of CO is 400 mm.   *Ans.* 10.05 cc

4. In a laboratory experiment 50 cc of a solution originally 0.2 molar in CH_3COOH was permitted to come to equilibrium with a 5 g sample of charcoal. At equilibrium 25 cc of the CH_3COOH solution required 30.0 cc of 0.1 N NaOH for neutralization. Calculate the weight of CH_3COOH adsorbed per gram of charcoal and the equilibrium concentration of the acid.

5. The following data were obtained for the adsorption of acetone on charcoal from an aqueous solution at 18° C.

y (millimoles/g)	C (millimoles/liter)
0.208	2.34
0.618	14.65
1.075	41.03
1.50	88.62
2.08	177.69
2.88	268.97

By plotting the proper functions, evaluate the constants k and n of the Freundlich equation.

6. For the adsorption of an organic dye on animal charcoal at 15° C the Freundlich constants are $n = 10.0$ and $k = 0.4$ where y is in milligrams per gram and C is in milligrams per liter. For the same dye at 15° C the Freundlich constants when blood charcoal is used as the adsorbent are $n = 21.0$ and $k = 0.7$ for the same units. Compare the effectiveness of the two charcoals per gram in removing the dye from a solution containing originally 1 mg/liter of dye. Use a method of successive approximations.

   Ans. Blood charcoal 1.68 times as effective as animal charcoal

7. A 0.1 molar solution of a certain soap has a surface tension 10 dynes/centimeter less than that of pure H_2O at 27° C. Estimate the excess concentration, in moles per square centimeter, of soap in the surface of the solution.

   Ans. 4.01×10^{-10} moles/sq cm

8. Using the data listed in Table 5, estimate the freezing point and osmotic pressure at 27° C of a solution containing 1 g of egg albumin per 100 g of H_2O.

9. A solution containing 0.2 mg/liter of a suspended material whose density is 2.2 g/cubic centimeter is observed under the ultramicroscope in a field of view 0.04 mm in diameter and 0.03 mm in depth. On the average the field of view was found to contain 8.5 particles. Assuming them to be spherical, what is the diameter of the particles?   *Ans.* 9.36×10^{-6} cm

10. The mean diameter of the particles in a certain aqueous colloidal solution is 42 Å. Calculate the diffusion coefficient at 25° C.

11. A portion of a certain suspension at 20° C requires 80 min to fall under the influence of gravity through a column of H_2O 1 m in height. If the density of the suspended particles is 2.0 g/cubic centimeter, what is the radius of the individual particles?

12. In measuring particle size in an aqueous suspension at 25° C by the sedimentation velocity method in a centrifugal field the boundary of the suspension moves from a position 10 cm from the axis of rotation to a position 16 cm away in a period of 30.0 min when the centrifuge performs 20,000 rpm. If the density of the suspended material is 1.25 g/cubic centimeter, what is the radius of the suspended particles?   *Ans.* 3.10×10^{-6} cm

13. The molecular weight of a protein was determined by sedimentation equilibrium in a centrifugal field. After equilibrium had been established, it was observed that the sol is 1.90 times as concentrated at a distance 10.0 cm from the axis of rotation as at a distance of 6.0 cm. The rotor is run at 6000 rpm, the temperature is 25° C, the density of the pure protein is 1.2 g/cubic centimeter, and that of the solvent is unity. From these data find the molecular weight of the protein.

Thermochemistry

Thermochemistry is the branch of physical chemistry that deals with the thermal changes accompanying chemical and physical transformations. Its aims are the determination of the quantities of energy emitted or absorbed as heat during the course of a transformation, as well as the development of methods of calculating these thermal readjustments without recourse to experiment.

The Importance of thermochemical information for the chemist and chemical engineer cannot be overestimated, especially when it is realized that all chemical and physical changes are accompanied by absorption or evolution of heat, and that these thermal changes enter intimately into many physicochemical calculations. From a purely practical point of view it is essential for the chemist or chemical engineer to know whether heat is evolved or absorbed in a particular reaction, and how much, for in one case provision must be made for removing that heat in order to effect the reaction, while in the second case provision must be made for supplying the necessary quantity of heat. Again, the heats of various metamorphoses in chemistry are required for many types of physicochemical calculations of both a practical and theoretical nature. Thus we have already seen how the heats of vaporization, sublimation, and fusion enter into the calculations involving the Clapeyron equation and the colligative properties of solutions. Further, we shall see later that such data are also necessary for calculations involving the variation of equilibrium constants with temperature, the entropy and free energy change, and other thermodynamic quantities. For these reasons it is imperative to consider some of the methods employed to determine heats of reaction experimentally, as well as some of the methods and principles that have been deduced from thermodynamics for evaluation of such thermal changes without appeal to experiment in each case.

The energy units in which thermal changes are usually expressed are the *calorie*, the *joule*, and the *large* or *kilocalorie*. A kilocalorie is equal to 1000 small calories.

Measurement of Thermal Changes. To determine directly the heat change involved in a reaction calorimeters are employed. A cal-

orimeter consists essentially of an insulated container filled with water in which is immersed the reaction chamber. In an exothermic reaction the heat generated is transferred to the water; the consequent temperature rise of the water is read from an accurate thermometer immersed in the water. Knowing the quantity of water present, its specific heat, and the change in temperature, the amount of heat evolved in the reaction may be calculated. Special corrections must be applied for radiation, rate of cooling of the calorimeter, temperature rise of the vessels, stirrer, etc. To avoid the latter corrections it is best to determine the heat capacity of the calorimeter by burning a definite amount of substance whose heat of combustion has been accurately measured. For this purpose, specially prepared samples of benzoic acid, naphthalene, or sugar may be used.

A similar apparatus may serve also with endothermic reactions, except that the temperature drop, instead of temperature increase, is measured.

An adiabatic bomb calorimeter used to determine heats of combustion of organic compounds is illustrated in Fig. 1. The bomb, or reaction chamber, is made of platinum-plated or corrosion-resistant alloy 'steel and consists of two cups joined by means of a heavy nut. B is a water jacket surrounding the bomb. Through the cover are inserted a precision thermometer C to measure the rise in the temperature of the water due to the reaction, and a mechanical stirrer D. A sample of the compound whose heat of combustion is to be determined is weighed into a nickel or platinum crucible suspended in the bomb, the upper cup is put on, and the bomb is filled with oxygen up to a pressure from 15 to 20 atmospheres. The bomb is then placed within the calorimeter, and the charge is ignited by passing an electric current through F, a weighed piece of fine iron wire embedded in the charge in the crucible. In calculating the heat of combustion, allowance must be made for the heat generated in burning this wire.

To avoid radiation corrections, the calorimeter proper is surrounded by another water jacket G which is at all times kept at the same temperature as the water within the calorimeter. This is accomplished by passing an electric current through the water. The electric input can be so regulated that the heating effect produced by the current is just enough to keep the temperatures within the two baths at all times the same. Under such conditions there can be no radiation from the inner bath. It is this latter arrangement which gives this type of apparatus the name of adiabatic (constant heat) calorimeter.

Heat of Reaction at Constant Volume and at Constant Pressure. There are two general conditions under which thermochemical measurements are made, namely, (1) constant volume, and (2) constant pressure. Under the first of these conditions the volume of the system

whose thermal change is sought is kept constant during the whole course of the measurement, and the entire process is conducted in a manner similar to the one described in the preceding section. In measurements at constant pressure, however, the system is kept either open to the
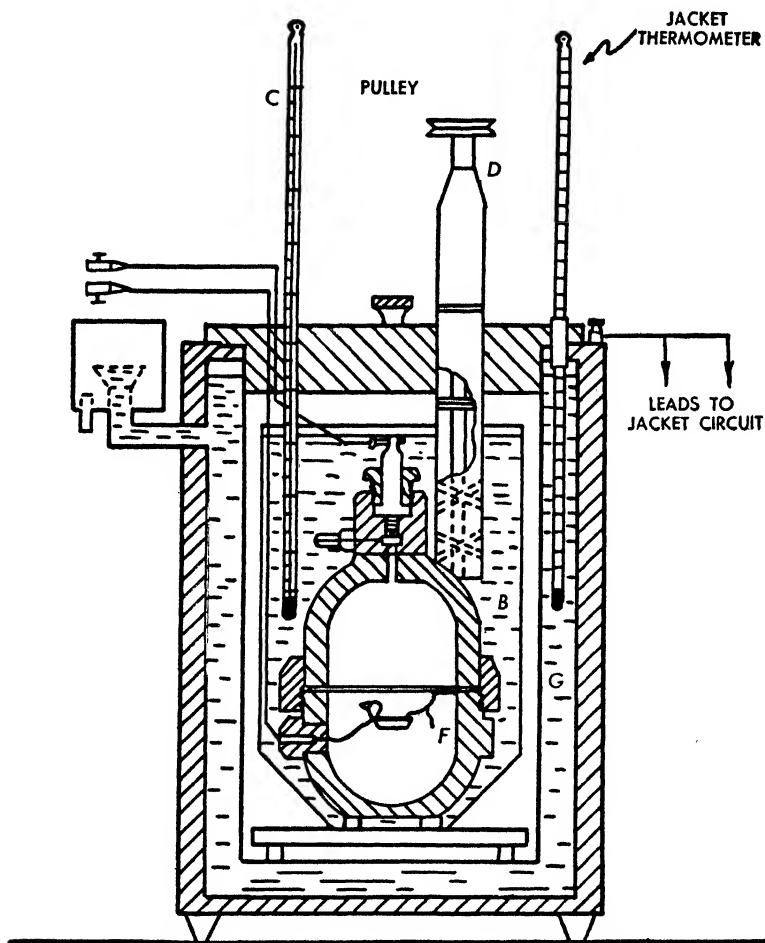
Fig. 1. Emerson Adiabatic Calorimeter

atmosphere, or is confined within a vessel upon which a constant external pressure is exerted. Under these conditions any volume change which accompanies the transformation or reaction can take place, and the system is able to adjust itself to the constant external pressure.

The magnitudes of the thermal changes obtained under these two conditions are, in general, different. At constant volume any thermal change taking place must be due only to the difference in the sum of the internal

energies of the products and the sum of the internal energies of the reactants. At constant pressure, however, not only does the change in the internal energy take place, but also either work is done on the surroundings, if there is an increase in volume during the change, or work is done by the surroundings on the system, if there is a diminution in volume on reaction. This work must modify the amount of heat observed in the calorimeter at constant volume.

The exact significance of both types of measurement can best be obtained by applying the first law of thermodynamics to the thermal change occurring within the calorimeter. According to the first law (see Chapter II), any heat q added to a system will in general go to increase the internal energy of the system and to perform external work, namely,

$$q = \Delta E + w = \Delta E + \int_{V_1}^{V_2} P dV \tag{1}$$

Since at constant volume $dV = 0$, no work can be performed. Hence, $\int_{V_1}^{V_2} P dV = 0$. At constant volume, therefore,

$$(q)_V = \Delta E \tag{2}$$

i.e., any thermal change occurring in the calorimeter must be the change in the internal energy due to the chemical reaction or physical transformation. On the other hand, when the pressure is kept constant,

$$w = \int_{V_1}^{V_2} P dV = P(V_2 - V_1) = P\Delta V$$

and, therefore, $\qquad (q)_P = \Delta E + P\Delta V \tag{3}$

According to equation (3) the thermal change observed in the calorimeter at constant pressure involves not only the change in the internal energy, but also the work performed in any expansion or contraction of the system.

Equation (3) may be rewritten to

$$\begin{aligned} (q)_P &= (E_p - E_r) + P(V_p - V_r) \\ &= (E_p + PV_p) - (E_r + PV_r) \end{aligned} \tag{4}$$

where the subscripts p and r refer to products and reactants respectively. But, since H, the heat content, is defined by $H = E + PV$, $H_p = E_p + PV_p$ and $H_r = E_r + PV_r$, and equation (4) becomes

$$\begin{aligned} (q)_P &= H_p - H_r \\ &= \Delta H \end{aligned} \tag{5}$$

We see, therefore, that whereas the thermal change at constant volume measures the change in the internal energy of products and reactants, the thermal change measured at constant pressure gives the change in

the heat contents of products and reactants. Equation (2) defines the *heat of reaction at constant volume* and shows that it is equal to the thermodynamic quantity ΔE. Similarly, equation (5) defines the *heat of reaction at constant pressure* and shows that it is equal to ΔH. The relation between the two follows, of course, from equations (3) and (5), namely,

$$(q)_P = \Delta E + P\Delta V = \Delta H$$
or $$\Delta H = \Delta E + P\Delta V \tag{6}$$

In line with the conventions established in Chapter II for ΔH and ΔE, a *positive value of ΔH or ΔE shows that heat is absorbed* during the process, while a *negative value of ΔH or ΔE shows that heat is evolved.*

Thermochemical Equations. The heat evolved or absorbed in a chemical or physical transformation depends not only on whether the change occurs at constant volume or constant pressure, but also on the amounts of substances considered, their physical state, the temperature, and the pressure. The effect of pressure on the heat of reaction is usually quite small; and, since considerations of the variation of heat of reaction with pressure involve a more complete knowledge of thermodynamics than the student possesses at present, they will be omitted in this discussion. The variation of heat of reaction with temperature will be discussed later in the chapter. At present we shall consider the effects on the heat of reaction of the quantities of substances involved and their physical state.

The actual amount of heat obtained in a reaction depends on the amount of substance reacted, and is directly proportional to it. Thus, when 2 grams of hydrogen are burned in oxygen to form liquid water, 68,400 calories of heat are evolved. For 4 grams the amount of heat evolved would be twice 68,400 calories, and so on. Instead of recording the amount of heat evolved per gram of a particular substance reacted, it is customary, rather, to give the heat for a particular reaction, namely,

$$H_2 + \frac{1}{2} O_2 = H_2O \qquad \Delta H = -\ 68{,}400 \text{ cal} \tag{7a}$$
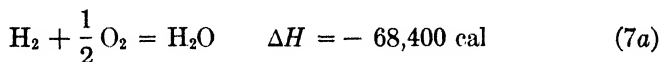
Equation ($7a$) is called a *thermochemical equation.* It indicates that when 1 mole of hydrogen (2.016 grams) reacts with 0.5 mole of oxygen (16 grams) to form 1 mole of water (18.016 grams) 68,400 calories of heat are evolved at constant pressure. If it is desired to indicate the interaction of 2 moles of hydrogen with oxygen to form 2 moles of water, then equation ($7a$) becomes

$$2\,H_2 + O_2 = 2\,H_2O \qquad \Delta H = 2(-\ 68{,}400) = -\ 136{,}800 \text{ cal} \tag{7b}$$
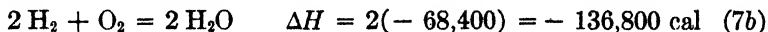
In writing a thermochemical equation it is essential that the equation be *balanced* and that the values of ΔE or ΔH indicated in the reaction correspond to the quantities of substance given by the equation.

Still, equation (7a) is incomplete in two respects. In the first place, the heat of reaction given corresponds to a definite temperature, in this instance 25° C, and this must be stated. Second, the equation as it stands does not tell us anything about the physical state of reactants and products, and yet such information is necessary to define the conditions for which ΔH is given. There is not much ambiguity about the state of hydrogen and oxygen, for both are definitely gaseous at this temperature, but the water may be liquid or gaseous. For the formation of liquid water from the elements at 25° C, $\Delta H = -68,400$ calories per mole. However, if a mole of gaseous water is formed from the elements at the same temperature, $\Delta H = -57,800$ calories per mole, the two values of ΔH differing from each other by the heat of vaporization of a mole of water at 25° C, namely, 10,600 calories. To make the thermochemical equation complete, therefore, we write, instead of equation (7a),

$$\text{H}_2(\text{g}) + \frac{1}{2}\,\text{O}_2(\text{g}) = \text{H}_2\text{O}(\text{l}) \qquad \Delta H_{25°\,\text{C}} = -68,400 \text{ cal} \qquad (8)$$
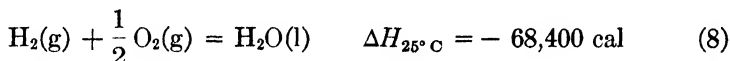
The symbols (g) and (l) indicate that the hydrogen and oxygen are gaseous, the water is liquid, while the subscript on the ΔH shows that the heat of reaction given is for 25° C or 298.2° K. In a similar manner the symbol (s) is used to designate a solid phase, while the symbol (aq) is used to represent a dilute aqueous solution.

It is very important that the student appreciate the physical significance of a thermochemical equation. Since in general ΔH is defined as $\Delta H = H_{\text{(products)}} - H_{\text{(reactants)}}$, ΔH for the reaction as given in equation (8) defines the difference in heat content of 1 mole of liquid water and the sum of the heat contents of 1 mole of gaseous hydrogen and 0.5 mole of gaseous oxygen, namely,

$$\Delta H = H_{\text{H}_2\text{O(l)}} - \left(H_{\text{H}_2(\text{g})} + \frac{1}{2} H_{\text{O}_2(\text{g})} \right)$$

As ΔH in equation (8) is negative, the sum of the heat contents of the reactants must exceed that of the product by 68,400 calories, and consequently when a mole of liquid water is formed from the elements that amount of heat is *evolved*. Conversely, if we were to consider the reverse of the reaction given, namely, the decomposition of a mole of water to form the elements, the sum of the heat contents of products would exceed the heat content of the reactant by the same amount, and 68,400 calories of heat would have to be *absorbed* to accomplish the reaction. It is evident, therefore, that once ΔH for a particular reaction is known, the ΔH for the reverse reaction is also known from the relation

$$\Delta H_{\text{direct reaction}} = -\Delta H_{\text{reverse reaction}} \qquad (9)$$

Calculation of ΔE from ΔH and Vice Versa. According to equation (6), the difference between ΔH and ΔE is given by $P\Delta V$, where P is constant. The significant factor in determining the difference $\Delta H - \Delta E$ is, therefore, ΔV, the change in volume which occurs during a reaction. For reactions involving solids and liquids only, the volume changes are in general very slight, and may be disregarded. For such reactions, then, $\Delta H - \Delta E = 0$, and ΔH is numerically equal to ΔE. In reactions involving gases, however, the volume changes may be large and cannot be disregarded. For such reactions the difference $\Delta H - \Delta E$ can be calculated very simply if we assume that the gases are ideal. If in general we have a reaction in which are involved n_r moles of *gaseous* reactants and n_p moles of *gaseous* products, then $PV_r = n_r RT$, $PV_p = n_p RT$, and, since the pressure is constant,

$$
\begin{aligned}
P\Delta V &= P(V_p - V_r) \\
&= n_p RT - n_r RT \\
&= (n_p - n_r)RT \\
&= \Delta n RT \qquad\qquad (10)
\end{aligned}
$$

where $\Delta n = (n_p - n_r)$, i.e., *the difference between the number of moles* of *gaseous products and reactants.* If equation (10) is inserted into equation (6), we obtain

$$\Delta H = \Delta E + \Delta n RT \qquad\qquad (11)$$

The usefulness of this equation in converting ΔH into ΔE, and vice versa, can best be illustrated by an example. For the combustion of 1 mole of liquid benzene, the heat of reaction at constant volume, ΔE, is given by

$$C_6H_6(l) + 7\tfrac{1}{2}\,O_2(g) = 6\,CO_2(g) + 3\,H_2O(l) \qquad \Delta E_{25^\circ C} = -\,781,100 \text{ cal}$$
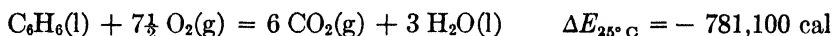
In this reaction there is a contraction in volume from 7.5 to 6 moles of gas, and hence $\Delta n = 6 - 7.5 = -1.5$. From equation (11) we obtain, then, for ΔH,

$$
\begin{aligned}
\Delta H_{25^\circ C} &= \Delta E_{25\,^\circ C} + \Delta n RT \\
&= -\,781,100 - 1.5(1.99)(298.2) \\
&= -\,781,100 - 894 \\
&= -\,782,000 \text{ cal}
\end{aligned}
$$

Hess's Law of Heat Summation. A very important principle of thermochemistry, based on the first law of thermodynamics but discovered independently, was enunciated by Hess in 1840. Hess showed that *the heat evolved or absorbed in a given reaction must always be constant and independent of the particular manner in which the reaction takes place.* It depends only on the initial and final states of the system and is not at

all affected by the number of steps which may intervene between the reactants and products. Stated differently, if a reaction proceeds in several steps, the heat of the overall reaction will be the algebraic sum of the heats of the various stages, and this sum in turn will be identical with the heat the reaction would evolve or absorb were it to proceed in a single step. This generalization is known as *Hess's law of constant heat summation*, or simply as *Hess's law*.

This principle makes it possible to calculate the heats of many reactions which either do not lend themselves to direct experimental determination or which it is not desired to measure. In these calculations thermochemical equations are handled as if they were ordinary algebraic equations, being added, subtracted, and multiplied or divided. The manner in which this is done can be illustrated by the following example. Suppose it is desired to find ΔH for the reaction

$$2\,C(s) + 2\,H_2(g) + O_2(g) = CH_3COOH(l) \qquad \Delta H_{25°\,C} = ? \qquad (12)$$

The heat of this reaction cannot be determined directly. However, calorimetric measurements are available for the reactions:

$$CH_3COOH(l) + 2\,O_2(g) = 2\,CO_2(g) + 2\,H_2O(l)$$
$$\Delta H_{25°\,C} = -\,207,900 \text{ cal} \quad (12a)$$

$$C(s) + O_2(g) = CO_2(g) \qquad\qquad \Delta H_{25°\,C} = -\,94,480 \text{ cal} \quad (12b)$$

$$H_2(g) = \frac{1}{2}\,O_2(g) = H_2O(l) \qquad\quad \Delta H_{25°\,C} = -\,68,400 \text{ cal} \quad (12c)$$
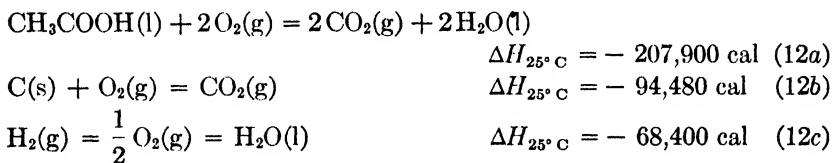
If we multiply now equations (12b) and (12c) by 2 and then add, we have:

$$2\,C(s) + 2\,O_2(g) = 2\,CO_2(g) \qquad \Delta H_{25°\,C} = -\,188,960 \text{ cal}$$
$$2\,H_2(g) + O_2(g) = 2\,H_2O(l) \qquad \Delta H_{25°\,C} = -\,136,800 \text{ cal}$$

$$2C(s) + 2H_2(g) + 3O_2(g) = 2CO_2(g) + 2H_2O(l)$$
$$\Delta H_{25°\,C} = -\,325,760 \text{ cal} \quad (12d)$$
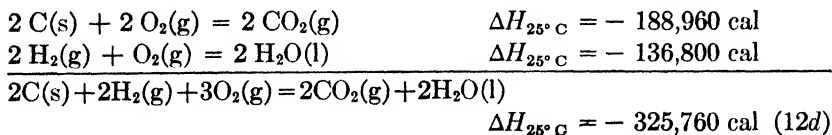
On subtracting (12a) from (12d), we obtain

$$2C(s) + 2H_2(g) + O_2(g) - CH_3COOH(l) = 0 \qquad \Delta H_{25°\,C} = -\,117,860 \text{ cal}$$

and, therefore,

$$2\,C(s) + 2\,\dot{H}_2(g) + O_2(g) = CH_3COOH(l) \qquad \Delta H_{25°\,C} = -\,117,860 \text{ cal}$$
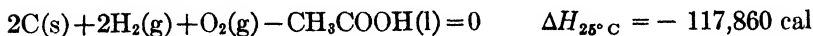
$\Delta H_{25°\,C} = -\,117,900$ calories is then the heat of reaction at constant pressure for equation (12).

Heats of Formation. A thermochemical equation of the type

$$\cdot\ \ Ag(s) + \frac{1}{2}\,Cl_2(g) = AgCl(s) \qquad \Delta H_{25°\,C} = -\,30,300 \text{ cal} \qquad (13)$$
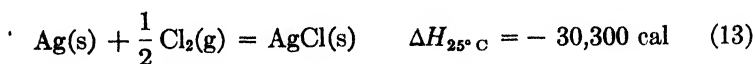
indicates that in the formation of 1 mole of silver chloride from the elements 30,300 calories of heat are evolved. Were the heat contents of the elements known, the above information would be sufficient to calculate the heat content of silver chloride. However, such information is at present unavailable. To get around this difficulty, the assumption is arbitrarily made that *the heat contents of all elements in the physical state and allotropic form stable under 1 atmosphere pressure are zero.* The temperature chosen is usually 18 or 25° C. On this basis the heat contents of $Ag(s)$ and $Cl_2(g)$ are both zero. Hence,

$$\Delta H_{25°} = -30{,}300 = H_{AgCl} - \left(H_{Ag} + \frac{1}{2}H_{Cl_2}\right)$$
$$= H_{AgCl} - 0$$
$$= H_{AgCl}$$

i.e., the thermal change involved in the formation of a compound from the *elements* may be considered to be the heat content of the compound. Such a thermal change is referred to as the *heat of formation* of the compound.

The heat of formation of a compound may be obtained either by measuring the ΔH of the reaction involving the direct formation of the compound from the elements, or by calculating the heat of formation from heats of reactions involving the compound. For instance, to find the heat of formation of $Fe_2O_3(s)$, the reactions

$$3\,C(s) + 2\,Fe_2O_3(s) = 4\,Fe(s) + 3\,CO_2(g) \quad \Delta H_{25°\,C} = +113{,}600 \text{ cal} \quad (14a)$$
$$C(s) + O_2(g) = CO_2(g) \qquad\qquad \Delta H'_{25°\,C} = -94{,}500 \text{ cal} \quad (14b)$$
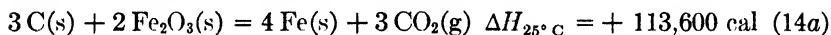
are available. In terms of the assumption made above, the heats of formation of $C(s)$, $Fe(s)$, and $O_2(g)$ are zero, and therefore,

$$\Delta H_{25°\,C} = +113{,}600 = 3\,\Delta H_{CO_2} - 2\,\Delta H_{Fe_2O_3}$$
$$\Delta H'_{25°\,C} = \Delta H_{CO_2} = -94{,}500 \text{ cal}$$

or

$$\Delta H_{Fe_2O_3} = \frac{3\,\Delta H_{CO_2} - \Delta H}{2}$$
$$= \frac{3(-94{,}500) - 113{,}600}{2}$$
$$= -198{,}600 \text{ cal per mole of } Fe_2O_3$$

The heats of formation for a number of compounds at 25° C are given in Table 1. With the aid of such a table heats of reaction may be easily calculated. Thus, for the reaction

$$Na_2CO_3(s) + 2\,HCl(g) = 2\,NaCl(s) + CO_2(g) + H_2O(l) \quad (15)$$

ΔH is given by

$$\Delta H_{25°\,C} = (2\,\Delta H_{NaCl} + \Delta H_{CO_2} + \Delta H_{H_2O}) - (\Delta H_{Na_2CO_3} + 2\,\Delta H_{HCl})$$

Substituting the appropriate heats of formation from the table, we obtain

$$\Delta H_{25°\,C} = [2(-98{,}300) + (-94{,}500) + (-68{,}400)]$$
$$- [(-271{,}000) + 2(-22{,}000)]$$
$$= -44{,}500 \text{ cal}$$

Hence ΔH of reaction (15) is $-44{,}500$ calories at 25° C.

Compounds formed from their elements with evolution of heat are known as exothermic compounds, while those formed with absorption of heat are referred to as endothermic compounds. The first type is much more common.

TABLE 1

HEATS OF FORMATION OF COMPOUNDS AT 25° C

Substance	ΔH cal/mole	Substance	ΔH cal/mole
$HCl(g)$	$-22{,}000$	$NaOH(s)$	$-102{,}000$
$KCl(s)$	$-104{,}400$	$KOH(s)$	$-102{,}000$
$NaCl(s)$	$-98{,}300$	$HNO_3(l)$	$-42{,}000$
$AgCl(s)$	$-30{,}300$	$KNO_3(s)$	$-118{,}100$
$HgCl_2(s)$	$-53{,}400$	$AgNO_3(s)$	$-29{,}400$
$CuCl_2(s)$	$-53{,}400$	$NaNO_3(s)$	$-111{,}700$
$MgCl_2(s)$	$-153{,}300$	$CuNO_3(s)$	$-73{,}100$
$ZnCl_2(s)$	$-99{,}500$	$H_2S(g)$	$-4{,}800$
$NaBr(s)$	$-86{,}700$	$CuS(s)$	$-11{,}600$
$HgBr_2(s)$	$-40{,}700$	$ZnS(s)$	$-41{,}500$
$KI(s)$	$-78{,}900$	$H_2SO_4(l)$	$-189{,}800$
$NaI(s)$	$-69{,}300$	$K_2SO_4(s)$	$-341{,}800$
$H_2O(l)$	$-68{,}400$	$Na_2SO_4(s)$	$-332{,}200$
$H_2O(g)$	$-57{,}800$	$CuSO_4(s)$	$-184{,}300$
$CO(g)$	$-26{,}600$	$MgSO_4(s)$	$-313{,}000$
$CO_2(g)$	$-94{,}480$	$NiSO_4(s)$	$-213{,}600$
$NO(g)$	$+21{,}500$	$CaCO_3(s)$	$-283{,}500$
$NO_2(g)$	$+7{,}900$	$Na_2CO_3(s)$	$-271{,}000$
$N_2O(g)$	$+19{,}700$	$NH_3(g)$	$-11{,}050$
$SO_2(g)$	$-70{,}900$	$CH_4(g)$	$-18{,}100$
$Ag_2O(s)$	$-6{,}900$	$C_6H_6(l)$	$+11{,}600$
$CuO(s)$	$-38{,}500$	$C_2H_2(g)$	$+53{,}900$
$HgO(s)$	$-21{,}800$	$CH_3OH(l)$	$-57{,}500$
$FeO(s)$	$-64{,}300$	$CH_3OH(g)$	$-48{,}400$
$Fe_2O_3(s)$	$-198{,}500$	$C_2H_5OH(l)$	$-67{,}100$
$BaO(s)$	$-126{,}100$	$C_2H_5OH(g)$	$-56{,}900$

Heats of Combustion. The heats evolved in the complete combustion of many organic compounds in oxygen have been carefully determined. The interest many research workers have shown in this direction may be traced to the comparative smoothness with which this type of reaction proceeds as well as to the valuable information such data yield. The method ordinarily used is to burn the substance in a combustion bomb and to measure the heat evolved in a manner identical to the one previously described. Since heats of combustion are usually obtained

at constant volume, equation (11) is employed to convert the experimentally measured ΔE to ΔH.

The term *heat of combustion* refers to the amount of heat liberated *per mole* of substance burned.

Heats of combustion may be employed directly to calculate heats of formation of organic compounds. If the organic compounds contain only carbon, hydrogen, and oxygen, the supplementary information required is the heats of formation of carbon dioxide and water, the usual final oxidation products of such compounds. The method of calculation may be illustrated with the data on the combustion of propane. For gaseous propane, the heat of combustion is

$$C_3H_8(g) + 5\ O_2(g) = 3\ CO_2(g) + 4\ H_2O(l) \qquad \Delta H_{25°\,C} = -530,600 \text{ cal} \quad (16)$$
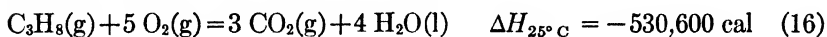
Therefore,

$$\Delta H_{25°\,C} = -530,600 = (3\ \Delta H_{CO_2} + 4\ \Delta H_{H_2O(l)}) - \Delta H_{propane}$$
$$\Delta H_{propane} = 530,600 + 3(-94,500) + 4(-68,400)$$
$$= -26,500 \text{ cal}$$

The heat of formation of propane from the elements is thus

$$3\ C(s) + 4\ H_2(g) = C_3H_8(g) \qquad \Delta H_{25°\,C} = -26,500 \text{ cal/mole} \quad (17)$$

Another direction in which heats of combustion have proved of value is in the study of the energy differences of allotropic forms of the elements. Carbon exists in two crystalline forms, diamond and graphite. When these two forms of carbon are burned, the heats evolved are found to be

$$C\ (\text{diamond}) + O_2(g) = CO_2(g) \qquad \Delta H_{25°\,C} = -94,480 \text{ cal} \quad (18)$$
$$C\ (\text{graphite}) + O_2(g) = CO_2(g) \qquad \Delta H_{25°\,C} = -94,030 \text{ cal} \quad (19)$$

Therefore, carbon in the diamond form has the higher heat content, and for the transition,

$$C\ (\text{graphite}) = C\ (\text{diamond}) \qquad \Delta H_{25°\,C} = 450 \text{ cal} \quad (20)$$

i.e., in the conversion of graphitic carbon into diamond, 450 calories of heat would be absorbed per gram atom.

Further information supplied by heats of combustion pertains to the energy associated with certain atomic groups in the molecule. It has been observed, for instance, that the heat of combustion in a homologous series of compounds varies by more or less constant amount as we pass from one member to the next higher one in the series. Thus for addition of a CH_2 group to a normal saturated paraffin hydrocarbon chain, the increase in ΔH of combustion is approximately 157,000 calories, as may be seen from Table 2. Similar regularities have been observed for other

groupings and linkages, and not only in the heats of combustion, but also in heats of formation. As may be anticipated, the thermal increments vary both with the nature of the group and the character of the bond.

TABLE 2

HEATS OF COMBUSTION OF HOMOLOGOUS PARAFFIN
HYDROCARBONS AT 25° C

Substance	ΔH cal/mole	Increase in ΔH per CH_2 Group
Methane (g)	− 212,800	
		160,000
Ethane (g)	− 372,800	
		157,800
Propane (g)	− 530,600	
		157,300
n-Butane (g)	− 687,900	
		157,400
n-Pentane (g)	− 845,300	

Heats of combustion of some compounds in addition to those already given are compiled in Table 3.

TABLE 3

HEATS OF COMBUSTION OF ORGANIC COMPOUNDS AT 25° C

Substance	Formula	ΔH (cal/mole)
Ethylene (g)	C_2H_4	− 337,300
Benzene (g)	C_6H_6	− 787,200
Benzene (l)	C_6H_6	− 782,000
Toluene (l)	C_7H_8	− 935,100
Naphthalene (s)	$C_{10}H_8$	− 1,231,800
Sucrose (s)	$C_{12}H_{22}O_{11}$	− 1,349,600
Methanol (l)	CH_3OH	− 173,600
Ethanol (l)	C_2H_5OH	− 326,700
Formic acid (l)	$HCOOH$	− 62,800
Acetic acid (l)	CH_3COOH	− 207,900
Oxalic acid (s)	$(COOH)_2$	− 60,100
Benzoic acid (s)	C_6H_5COOH	− 771,400
Salicylic acid (s)	HOC_6H_4COOH	− 723,100

Various empirical rules have been suggested for the calculation of heats of combustion. Thornton's [1] rule, applicable only to saturated paraffin hydrocarbons, states that ΔE of combustion per mole is given approximately by

$$\Delta E = (- 52,700)n \qquad (21)$$

where n is the number of *atoms* of oxygen required for the complete oxidation of 1 mole of paraffin hydrocarbon. Applied to propane, for instance, this rule would predict for the heat of combustion at constant volume $\Delta E = - 527,000$ calories per mole, as against the experimentally observed value of $\Delta E = - 528,800$ calories.

[1] Thornton, Phil. Mag., **33**, 196 (1917).

A much more elaborate and complete method of estimating the heats of combustion of various types of organic compounds has been suggested by Kharasch.[1] The student is referred to the original paper both for the details and for a very complete summary of heats of combustion.

Thermoneutrality of Salt Solutions. Since salts of strong acids and bases are considered to be completely ionized in dilute solutions, it might be expected that, if solutions of such salts were mixed without the incidence of some kind of chemical action, the heat effect observed should be essentially zero. Such is actually the case. Thus, when a dilute solution of potassium nitrate is mixed with a dilute solution of sodium bromide no heat effect is observed, namely,

$$KNO_3(aq) + NaBr(aq) = KBr(aq) + NaNO_3(aq) \qquad \Delta H = 0 \quad (22)$$
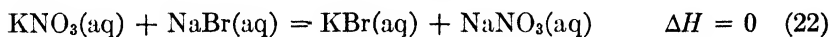
That the exchange indicated in equation (22) involves no chemical reaction may be seen more readily when the equation is written in ionic form:

$$K^+(aq) + NO_3^-(aq) + Na^+(aq) + Br^-(aq) =$$
$$K^+(aq) + Br^-(aq) + Na^+(aq) + NO_3^-(aq) \qquad \Delta H = 0 \quad (23)$$
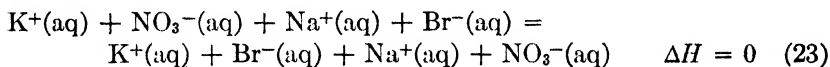
Since the products and reactants are identical, no thermal change is to be expected. If a slight thermal effect is observed, it is to be ascribed to the change in interionic attractions produced by the dilution on mixing.

The principle that dilute solutions of neutral salts of strong acids and strong bases may be mixed without absorption or evolution of heat is termed the *principle of thermoneutrality of salt solutions.*

When a chemical reaction does occur on mixing, the principle of thermoneutrality is no longer valid. Thus, when a dilute solution of barium chloride is mixed with a dilute solution of sodium sulfate, barium sulfate precipitates, and the heat of reaction instead of being zero is actually

$$BaCl_2(aq) + Na_2SO_4(aq) = BaSO_4(s) + 2 NaCl(aq)$$
$$\Delta H_{20° C} = - 5800 \text{ cal} \quad (24a)$$

or, in ionic form,

$$Ba^{++}(aq) + SO_4^{--}(aq) = BaSO_4(s) \qquad \Delta H_{20° C} = - 5800 \text{ cal} \quad (24b)$$
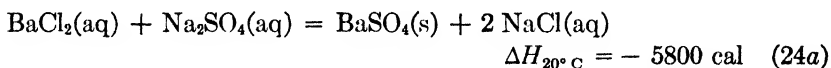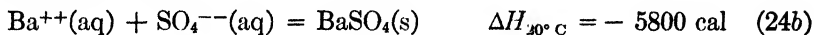
Heats of Neutralization of Acids and Bases. Hess and others have shown experimentally that, when dilute solutions of strong acids are neutralized with dilute solutions of strong bases, the *heat of neutralization* per mole of water formed is essentially constant and independent of the nature of the acid or base, as may be seen from the following data:

$$HCl(aq) + NaOH(aq) = NaCl(aq) + H_2O(l) \qquad \Delta H = - 13,680 \text{ cal}$$
$$HCl(aq) + LiOH(aq) = LiCl(aq) + H_2O(l) \qquad \Delta H = - 13,700 \text{ cal}$$
$$HNO_3(aq) + KOH(aq) = KNO_3(aq) + H_2O(l) \qquad \Delta H = - 13,870 \text{ cal}$$
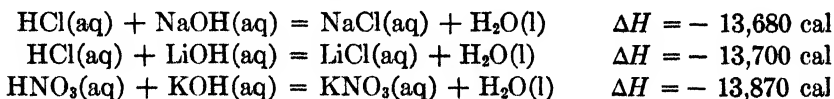
[1] Kharasch, Bureau of Standards Journal of Research, **2**, 359–430 (1929).

This constancy of the heat of neutralization is readily understood when it is remembered that strong acids, bases, and salts are completely dissociated in their dilute solutions, and that consequently the neutralization process involves only the combination of hydrogen and hydroxyl ions to form unionized water. Since this process is the same in all neutralizations, ΔH of neutralization should be essentially constant per mole of water formed. The value generally accepted at present for this thermal quantity is $-13,800$ calories at $20°$ C. Hence we may write, for the process,

$$H^+(aq) + OH^-(aq) = H_2O(l) \qquad \Delta H_{20° C} = -13,800 \text{ cal} \qquad (25)$$

Equation (25) represents the heat of formation of a mole of liquid water from the two ions, H^+ and OH^-.

TABLE 4

HEATS OF NEUTRALIZATION OF ACIDS AND BASES IN DILUTE SOLUTIONS
(room temperature)

Acid	Base	ΔH (cal)
HCl	NaOH	$-13,680$
HCl	NH$_4$OH	$-12,400$
HC$_2$H$_3$O$_2$	NaOH	$-13,300$
HC$_2$H$_3$O$_2$	NH$_4$OH	$-12,000$
HCOOH	NaOH	$-13,400$
HCOOH	NH$_4$OH	$-11,900$
1/2 H$_2$S	NaOH	$-3,800$
1/2 H$_2$S	NH$_4$OH	$-3,100$
HCN	NaOH	$-2,900$
HCN	NH$_4$OH	$-1,300$

This constancy of heat of neutralization does not carry over to the neutralization of weak acids by strong bases, weak bases by strong acids, or weak acids by weak bases, as may be seen from Table 4. The behavior of the reactions mentioned is explicable on the ground that in such neutralizations the combination of hydrogen and hydroxyl ions to form water is not the only reaction taking place. Take the case of hydrocyanic acid and sodium hydroxide as an example. In water solution the hydrocyanic acid is practically unionized. Before the hydrogen ion of the acid can react with the hydroxyl ion of the base ionization must take place. Since this ionization occurs while the neutralization is proceeding, the thermal change observed is the sum of the heat of ionization of the acid and the heat of neutralization of the ionized hydrogen ion; i.e., the overall reaction,

$$HCN(aq) + NaOH(aq) = NaCN(aq) + H_2O(l)$$
$$\Delta H = -2900 \text{ cal} \qquad (26a)$$
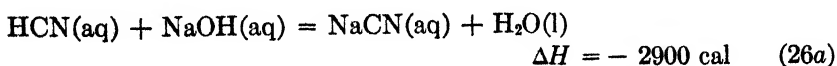
is in reality composed of the two reactions:

$$HCN(aq) = H^+(aq) + CN^-(aq) \qquad \Delta H = \Delta H_i \qquad (26b)$$
$$H^+(aq) + NaOH(aq) = Na^+(aq) + H_2O(l) \quad \Delta H = -13,800 \text{ cal} \quad (26c)$$
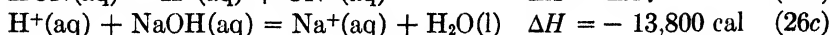
The sum of equations (26b) and (26c) gives equation (26a), and consequently,

$$\Delta H_i + (-13,800) = -2900$$
$$\Delta H_i = +10,900 \text{ cal} \qquad (27)$$

ΔH_i is the heat of ionization of the hydrocyanic acid per mole. For the purpose at hand, the slight ionization in water of an acid as weak as hydrocyanic acid may be disregarded.

In a similar manner may be explained the results on the neutralization of weak bases by strong acids, and weak bases by weak acids. In the first instance the heat of ionization of the weak base must be considered, while in the latter the heats of ionization of both the weak acid and the weak base come into play.

The Heats of Formation of Ions. For the calculation of heats of reactions involving electrolytes in aqueous solution a knowledge of the heats of formation of the electrolytes in solution from the elements is required. Such data can be obtained from a combination of the heat of formation of the pure substance from the elements, and the heat of solution of the substance in the solvent. The sum of these two heats is the heat of formation of the substance in solution. Thus, for instance, the heat of formation of calcium chloride(s) from the elements at 18° C is $\Delta H = -190,500$ calories per mole, while the heat of solution of a mole of calcium chloride(s) in sufficient water to form a dilute solution at the same temperature is $\Delta H = -18,000$ calories. Consequently,

$$Ca(s) + Cl_2(g) = CaCl_2(s) \qquad \Delta H_{18° C} = -190,500 \text{ cal} \quad (28a)$$
$$CaCl_2(s) + aq = CaCl_2(aq) \qquad \Delta H_{18° C} = -18,000 \text{ cal} \quad (28b)$$

$$Ca(s) + Cl_2(g) + aq = CaCl_2(aq) \qquad \Delta H_{18° C} = -208,500 \text{ cal} \quad (28c)$$
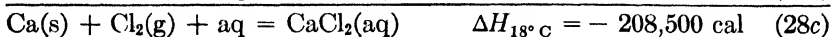
and this is the heat of formation of calcium chloride(aq) per mole.

Since the principle of thermoneutrality applies to strong electrolytes in dilute solutions, or, stated differently, since individual ions in dilute solutions may be considered to have thermal properties essentially independent of the accompanying ions, it would be more convenient to tabulate the heats of formation of the individual ions, if these could be calculated, rather than the heats of formation for the substance. Then, for any given electrolyte the heat of formation would be the sum of the heats of formation of the individual ions.

The approach to this problem can be made through equation (25). Since the heat of formation of a mole of water from the hydrogen and

hydroxyl ions involves the evolution of 13,800 calories of heat, this amount of heat must be supplied in order to dissociate a mole of water into these two ions, namely,

$$H_2O(l) = H^+(aq) + OH^-(aq) \qquad \Delta H = 13,800 \text{ cal} \qquad (29)$$
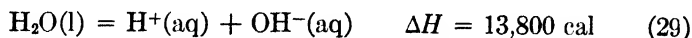
If equation (29) is combined with the equation for the heat of formation of 1 mole of $H_2O(l)$ from the elements,

$$H_2(g) + \frac{1}{2} O_2(g) = H_2O(l) \qquad \qquad \Delta H = -68,400 \text{ cal} \quad (30)$$

we obtain the heat of formation of the hydrogen and hydroxyl ions from the elements, or

$$H_2(g) + \frac{1}{2} O_2(g) = H^+(aq) + OH^-(aq) \qquad \Delta H = -54,600 \text{ cal} \quad (31)$$
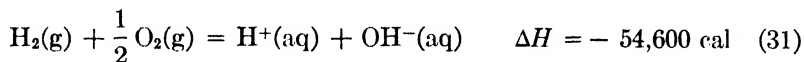
This sum cannot be resolved at present without some assumption with respect to the heat of formation of one of these ions. The convention generally adopted is that *the heat of formation of the hydrogen ion in solution is zero,* i.e., that for the reaction

$$\frac{1}{2} H_2(g) + aq = H^+(aq) \qquad \qquad \Delta H = 0 \qquad (32)$$

With this convention equation (31) gives directly the heat of formation of the hydroxyl ion, or

$$\frac{1}{2} H_2(g) + \frac{1}{2} O_2(g) + aq = OH^-(aq) \qquad \Delta H = -54,600 \text{ cal} \quad (33)$$
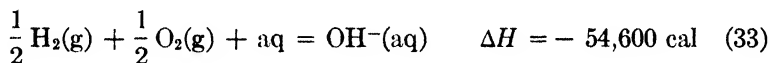
Once the heats of formation of these two ions are known, those for other ions may be readily calculated. Since the heat of formation of hydrochloric acid(aq), $(H^+ + Cl^-)$, from the elements at 18° C is $\Delta H = -39,550$, and since $\Delta H = 0$ for $H^+(aq)$, $\Delta H = -39,550$ must represent the heat of formation of the chloride ion. Again, in the reaction

$$Na(s) + \frac{1}{2} O_2(g) + \frac{1}{2} H_2(g) + aq = NaOH(aq) \quad \Delta H = -112,100 \text{ cal} \quad (34)$$
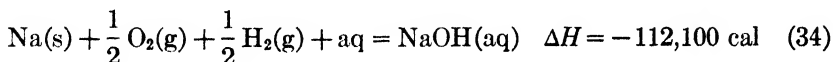
Combining this equation with (33), we see that

$$-112,100 = \Delta H_{Na^+} + \Delta H_{OH^-}$$
$$= \Delta H_{Na^+} - 54,600$$
$$\Delta H_{Na^+} = -57,500 \text{ cal}$$

By similar procedures the heats of formation of many other ions in aqueous solution have been evaluated. Some of these are summarized

in Table 5. It will be observed that the formation of most ions from the elements involves the evolution of large amounts of heat, some of it due, no doubt, to the hydration of the ions by the solvent.

TABLE 5

HEATS OF FORMATION OF IONS AT 20° C

Ion	ΔH (cal)	Ion	ΔH (cal)
H^+	0	Cd^{++}	$- 17,400$
Li^+	$- 66,500$	Hg^{++}	$+ 41,600$
Na^+	$- 57,500$	Sn^{++}	$- 2,400$
K^+	$- 60,300$	Pb^{++}	$- 500$
NH_4^+	$- 31,700$	OH^-	$- 54,600$
Ag^+	$+ 24,900$	Cl^-	$- 39,600$
Cu^+	$+ 16,000$	Br^-	$- 28,600$
Hg^+	$+ 20,100$	I^-	$- 13,300$
Mg^{++}	$- 110,200$	NO_2^-	$- 25,600$
Ca^{++}	$- 129,500$	NO_3^-	$- 49,800$
Sr^{++}	$- 130,200$	ClO_3^-	$- 19,100$
Ba^{++}	$- 128,300$	$C_2H_3O_2^-$	$- 117,600$
Mn^{++}	$- 49,700$	S^{--}	$+ 10,000$
Fe^{++}	$- 20,800$	SO_3^{--}	$- 146,800$
Co^{++}	$- 16,500$	SO_4^{--}	$- 211,200$
Ni^{++}	$- 15,300$	CO_3^{--}	$- 161,000$
Cu^{++}	$+ 16,500$	PO_4^{---}	$- 298,500$
Zn^{++}	$- 36,600$		

The use of Table 5 in the calculation of heats of reactions involving strong acids, bases, or salts in dilute aqueous solutions can best be shown by an example. Suppose it is desired to calculate ΔH for the reaction

$$\frac{1}{2} H_2(g) + AgCl(s) = Ag(s) + H^+(aq) + Cl^-(aq) \qquad \Delta H_{20° C} = ? \quad (35)$$
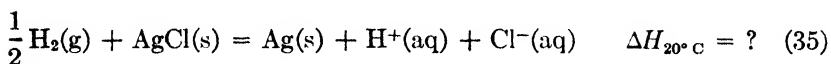
Since the heats of formation of $H_2(g)$, $Ag(s)$, and $H^+(aq)$ are all zero, ΔH is given by

$$\Delta H = \Delta H_{Cl^-} - \Delta H_{AgCl(s)}$$

On inserting the value of $\Delta H_{Cl^-} = - 39,600$ from Table 5, and, neglecting the small temperature difference, $\Delta H_{AgCl(s)} = - 30,300$ from Table 1, we get

$$\Delta H_{20° C} = - 39,600 - (- 30,300)$$
$$= - 9300 \text{ cal}$$

Integral Heats of Solution and Dilution. Solution of a substance in a solvent is accompanied by absorption or evolution of heat, and this thermal effect is termed the *integral heat of solution* of the substance. Per mole of dissolved substance the integral heat of solution depends

upon the amount of solvent in which solution takes place, as may be seen from Table 6. For this reason it is essential to specify the number of moles of solvent per mole of solute in giving a heat of solution, as in the following:

$$H_2SO_4(l) + 9\ H_2O(l) = H_2SO_4(9\ H_2O) \qquad \Delta H_{18^\circ\,C} = -\ 15,110\ \text{cal} \quad (36)$$
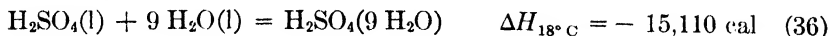
TABLE 6

INTEGRAL HEATS OF SOLUTION OF ONE MOLE OF H_2SO_4 IN WATER AT 18° C

Moles of Water (n_1)	ΔH (cal)
0.00	0
0.11	− 920
0.25	− 1,970
0.43	− 3,300
0.67	− 4,890
1.00	− 6,740
1.50	− 8,630
2.33	− 10,680
4.00	− 13,010
9.00	− 15,110
19.00	− 16,900
∞	− 20,200

However, when the amount of water per mole of substance is large, it is usually found that still larger amounts of solvent will produce no further sensible thermal effect. Once this state of a dilute solution has been reached, the symbol *aq* is employed, designating that further dilution will produce no additional thermal effect. Thus the limiting value of the integral heats of solution in Table 6 would be represented by

$$H_2SO_4(l) + aq = H_2SO_4(aq) \qquad \Delta H = -\ 20,200\ \text{cal} \qquad (37)$$

The difference between any two values of the integral heats of solution of a substance gives the heat which is involved in the *dilution* of the substance from the initial state to the final state, and is termed the *heat of dilution* of the substance. According to Table 6 the heat recoverable on diluting with 8 moles of water a solution containing 1 mole of sulfuric acid in 1 mole of water is

$$H_2SO_4(1\ H_2O) + 8\ H_2O = H_2SO_4(9\ H_2O)$$
$$\Delta H = -\ 15,110 - (-\ 6740) = -\ 8370\ \text{cal} \quad (38)$$
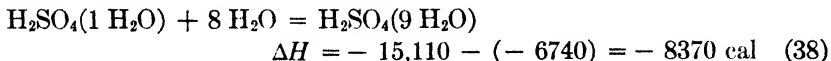
Similarly, the heat available on diluting the same solution with a very large quantity of water is

$$H_2SO_4(1\ H_2O) + aq = H_2SO_4(aq)$$
$$\Delta H = -\ 20,200 + 6740 = -\ 13,460\ \text{cal} \quad (39)$$
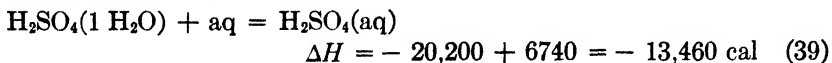
The latter value represents the maximum heat obtainable on diluting the given solution.

Heats of Hydration. Attention has already been directed to the utilization of integral heats of solution for calculating heats of formation in solution and heats of dilution. Integral heats of solution may also be employed to calculate *heats of hydration;* i.e., the heats required for reactions of the type

$$CaCl_2(s) + 2 H_2O(l) = CaCl_2 \cdot 2 H_2O(s) \qquad (40)$$
$$CaCl_2 \cdot 2 H_2O(s) + 2 H_2O(l) = CaCl_2 \cdot 4 H_2O(s) \qquad (41)$$
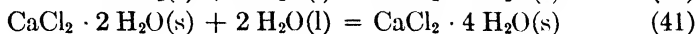
The integral heats of solution at 18° C for the various solid modifications of calcium chloride in the indicated quantities of water are given by

$$CaCl_2(s) + 400 H_2O(l) = CaCl_2(400 H_2O) \qquad \Delta H = -17,990 \text{ cal} \quad (42)$$
$$CaCl_2 \cdot 2 H_2O(s) + 398 H_2O(l) = CaCl_2(400 H_2O)$$
$$\Delta H = -10,030 \text{ cal} \quad (43)$$
$$CaCl_2 \cdot 4 H_2O(s) + 396 H_2O(l) = CaCl_2(400 H_2O)$$
$$\Delta H = -1830 \text{ cal} \quad (44)$$
$$CaCl_2 \cdot 6 H_2O(s) + 394 H_2O(l) = CaCl_2(400 H_2O)$$
$$\Delta H = +4560 \text{ cal} \quad (45)$$
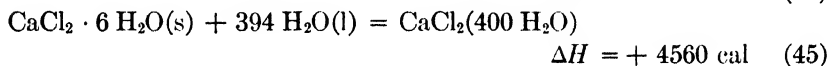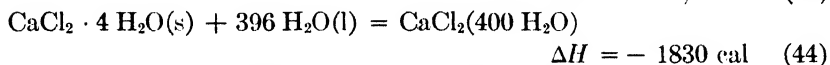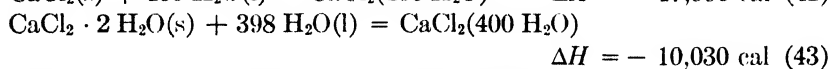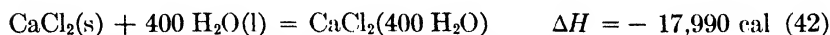
It will be noticed that all the final solutions formed are the same. Consequently the difference between any two of the equations (42) to (45) gives directly the *heat of hydration* of one solid phase to another. Thus, equation (42) minus equation (43) yields

$$CaCl_2(s) + 2 H_2O(l) = CaCl_2 \cdot 2 H_2O(s) \qquad \Delta H = -7960 \text{ cal} \quad (46)$$

Similarly (43) − (44) and (44) − (45) give

$$CaCl_2 \cdot 2 H_2O(s) + 2 H_2O(l) = CaCl_2 \cdot 4 H_2O(s) \quad \Delta H = -8200 \text{ cal} \quad (47)$$
$$CaCl_2 \cdot 4 H_2O(s) + 2 H_2O(l) = CaCl_2 \cdot 6 H_2O(s) \quad \Delta H = -6390 \text{ cal} \quad (48)$$

On the other hand, the heat evolved in the direct hydrations of $CaCl_2(s)$ to $CaCl_2 \cdot 6 H_2O(s)$ follows from the difference between equations (42) and (45), namely,

$$CaCl_2(s) + 6 H_2O(l) = CaCl_2 \cdot 6 H_2O(s) \qquad \Delta H = -22,500 \text{ cal} \quad (49)$$

Equations (42) to (45) emphasize the influence of the nature of the solid phase on the heat of solution. Whereas a mole of anhydrous calcium chloride will dissolve in water with the *evolution* of a large quantity of heat, dissolution of a mole of $CaCl_2 \cdot 6 H_2O$ can take place only with the *absorption* of 4560 calories of heat per mole of salt. The difference in the two heats of solution is due, of course, to the heat of hydration of $CaCl_2(s)$, and is given by equation (49).

Other Heats of Reaction. There are many types of heats of reaction besides those mentioned. Some of these have been discussed in other chapters, as heats of fusion, vaporization, and sublimation. Others, like

heats of dissociation, hydrogenation, polymerization, and adsorption, either require no further discussion or will be elaborated on in subsequent chapters. All these designations denote merely special names for particular types of heats of reaction. All of these heats are subject to the same treatment and thermodynamic requirements as the heats described in this chapter.

Variation of Heat of Reaction with Temperature. The heats of reaction obtained calorimetrically, by calculation from heats of formation, or by the methods to be outlined later apply to some one definite temperature. At other temperatures the heat of a reaction may be essentially the same or it may not, depending on the magnitude of the heat capacities of the reactants and products as well as on their variation with temperature. Generally it is not the same. As reactions are carried out at various temperatures, it is very frequently necessary to know heats of reaction at temperatures other than those at which they were determined. A method of calculating the heat of reaction at one temperature from that at another is, therefore, highly desirable.

In such calculations heat capacities play a very important part, and hence a word about these is in order. We have seen that there are two 'types of heat capacities, C_p and C_v. For gases these differ from each other by R, or $C_p - C_v = R$ calories per mole. On the other hand, since the volume changes accompanying the heating of solids and liquids are small, these two heat capacities may be taken as essentially equal for these. Furthermore, heat capacities, like heats of reaction, vary with temperature. This variation is frequently expressed by empirical formulas, as for instance,

$$CO_2(g): \quad C_p = 7.0 + 0.0071\ T - 0.00000186\ T^2 \quad (50a)$$
$$C(s, graphite): \quad C_p = 1.22 + 0.00489\ T - 0.00000111\ T^2 \quad (50b)$$

The use of these equations is limited to the range of temperatures over which they have been determined.

When a reaction,

$$Reactants\ (A) = Products\ (B) \quad (51)$$

taking place at a temperature T_1 and with a heat of reaction ΔH_1, is allowed to proceed at some higher temperature T_2, the thermal changes involved may be shown by the schematic diagram in Fig. 2. $A \longrightarrow B$ represents the reaction at T_1, $A' \longrightarrow B'$ the reaction at T_2. In the latter instance, reactants (A') react to give products (B') both at T_2 with

Fig. 2. Variation of Heat of Reaction with Temperature

a heat of reaction ΔH_2. If we were to determine, then, the heat contents of the reactants and products at T_2, the heat of reaction could be calculated.

Let H_A and H_B be the respective heat contents of A and B at T_1. To raise A from T_1 to T_2 heat equivalent to $C_{p_A}(T_2 - T_1)$, where C_{p_A} is the mean heat capacity of A between T_1 and T_2, must be supplied. The heat content of the reactants at T_2 is, therefore,

$$H'_A = H_A + C_{p_A}(T_2 - T_1) \tag{52}$$

Similarly the heat content of the products at T_2 is

$$H'_B = H_B + C_{p_B}(T_2 - T_1) \tag{53}$$

ΔH_2 at T_2 follows, then, as

$$\begin{aligned}\Delta H_2 = H'_B - H'_A &= H_B + C_{p_B}(T_2 - T_1) - H_A - C_{p_A}(T_2 - T_1) \\ &= (H_B - H_A) + (C_{p_B} - C_{p_A})(T_2 - T_1)\end{aligned} \tag{54}$$

But $H_B - H_A = \Delta H_1$; and, letting $(C_{p_B} - C_{p_A})$, *the heat capacity of the products minus the heat capacity of the reactants*, be ΔC_p, we have

$$\Delta H_2 = \Delta H_1 + \Delta C_p(T_2 - T_1) \tag{55}$$

The heat of reaction at any temperature is, therefore, the sum of the heat of reaction at the first temperature plus the difference in heat capacities of products and reactants multiplied by the temperature difference. This equation is strictly applicable only where ΔC_p is independent of temperature, or where the two temperatures are close enough so that there can be no great variation in C_p and constancy can be assumed.

To illustrate the use of equation (55) suppose it is desired to calculate ΔH at 348° K for the reaction

$$\frac{1}{2}\,H_2(g) + \frac{1}{2}\,Cl_2(g) = HCl(g) \quad \Delta H_{288°\,K} = -22{,}000 \text{ cal} \tag{56}$$
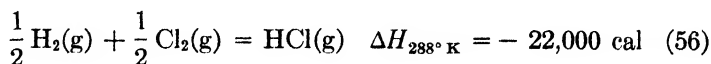
The mean specific heats are

$$\begin{aligned} H_2(g)\text{:} \quad & C_p = 6.82 \text{ cal/mole} \\ Cl_2(g)\text{:} \quad & C_p = 7.71 \text{ cal/mole} \\ HCl(g)\text{:} \quad & C_p = 6.81 \text{ cal/mole} \end{aligned}$$

from which $\Delta C_p = 6.81 - \dfrac{1}{2}(6.82) - \dfrac{1}{2}(7.71) = -0.46$

Therefore, $\Delta H_{348°\,K} = -22{,}000 + (-0.46)(348 - 288)$
$$= -22{,}028 \text{ cal}$$

Heat of Reaction Equations. When ΔC_p is not constant integration has to be resorted to to obtain ΔH. The differential equation for the variation of heat of reaction with temperature may be obtained as follows. Since

$$\Delta H = H_{\text{products}} - H_{\text{reactants}} \tag{57}$$

differentiation of both sides with respect to absolute temperature at constant pressure yields

$$\left[\frac{\partial(\Delta H)}{\partial T}\right]_P = \left[\frac{\partial H_{\text{products}}}{\partial T}\right]_P - \left[\frac{\partial H_{\text{reactants}}}{\partial T}\right]_P$$

But the second and third terms are respectively $C_{p\text{products}}$ and $C_{p\text{reactants}}$. Therefore,

$$\left[\frac{\partial(\Delta H)}{\partial T}\right]_P = C_{p\text{products}} - C_{p\text{reactants}} = \Delta C_p \qquad (58)$$

Equation (58) is known as *Kirchhoff's equation.* The calculation of ΔH from this relation can be illustrated by a specific example.

Suppose it is desired to find the heat of formation of ammonia at 1000° K. For the purpose at hand the following data are available

$$\frac{1}{2}\,N_2(g) + \frac{3}{2}\,H_2(g) = NH_3(g) \qquad \Delta H_{298.2°\,K} = -11{,}050 \text{ cal} \quad (59)$$

$$N_2(g): \quad C_p = 6.50 + 1.0 \times 10^{-3}\,T \text{ cal/mole} \qquad (60a)$$

$$H_2(g): \quad C_p = 6.62 + 0.81 \times 10^{-3}\,T \text{ cal/mole} \qquad (60b)$$

$$NH_3(g): \quad C_p = 7.12 + 6.09 \times 10^{-3}\,T - \frac{0.398 \times 10^5}{T^2} \text{ cal/mole} \quad (60c)$$

Now $C_{p\text{products}} = C_{p\text{NH}_3}$. Again, $C_{p\text{reactants}}$ is given by

$$\begin{aligned}
C_{p\text{reactants}} &= \frac{1}{2}\,C_{p\text{N}_2} + \frac{3}{2}\,C_{p\text{H}_2} \\
&= 3.25 + 0.5 \times 10^{-3}\,T + 9.93 + 1.22 \times 10^{-3}\,T \\
&= 13.18 + 1.72 \times 10^{-3}\,T \qquad (60d)
\end{aligned}$$

Subtracting equation (60d) from equation (60c), we obtain for ΔC_p

$$\Delta C_p = -6.06 + 4.37 \times 10^{-3}\,T - \frac{0.398 \times 10^5}{T^2} \qquad (60e)$$

On substitution of equation (60e) for ΔC_p in equation (58), the latter becomes

$$\left[\frac{\partial(\Delta H)}{\partial T}\right]_P = \Delta C_p = -6.06 + 4.37 \times 10^{-3}\,T - \frac{0.398 \times 10^5}{T^2}$$

and hence, on integration,

$$\begin{aligned}
\Delta H &= \int \Delta C_p\,dT + \Delta H_0 \\
&= \int \left(-6.06 + 4.37 \times 10^{-3}\,T - \frac{0.398 \times 10^5}{T^2}\right) dT + \Delta H_0 \\
&= -6.06\,T + 2.19 \times 10^{-3}\,T^2 + \frac{0.398 \times 10^5}{T} + \Delta H_0 \qquad (61)
\end{aligned}$$

ΔH_0 is the constant of integration. In order to evaluate this constant a known value of ΔH at some definite temperature must be substituted in equation (61), and ΔH_0 solved for. Since we have given that at 298.2° K $\Delta H = -11,050$ calories, then

$$- 11,050 = - 6.06(298.2) + 2.19 \times 10^{-3}(298.2)^2 + \frac{0.398 \times 10^5}{298.2} + \Delta H_0$$

and $\Delta H_0 = -11,050 + 1810 - 190 - 130$
$$= -9560 \text{ cal}$$

Inserting the found value of ΔH_0 into equation (61), the expression for ΔH becomes

$$\Delta H = -9560 - 6.06\,T + 2.19 \times 10^{-3}\,T^2 + \frac{0.398 \times 10^5}{T} \quad (62)$$

This equation gives ΔH for the synthesis of ammonia as a function of the temperature. Equations of this type are very useful, for they permit the calculation of the heat of reaction at any desired temperature merely by inserting the temperature in degrees absolute at which ΔH is sought. Thus for the present case of ΔH at 1000° K we have

$$\Delta H = -9560 - 6.06(1000) + 2.19 \times 10^{-3}(1000)^2 + \frac{0.398 \times 10^5}{1000}$$
$$= -9560 - 6060 + 2190 + 40$$
$$= -13,390 \text{ cal at } 1000° \text{ K}$$

Hence 2340 more calories of heat are evolved at 1000° K than at 298° K.

The accuracy of these equations is determined by the precision with which the specific heats and the heats of reaction have been determined. Discretion should be exercised as to the temperature range over which these equations are to be used, since they are valid only through the temperature interval over which the specific heats have been determined.

Reference Literature on Thermochemical Data. A student of physical chemistry should be familiar with the various publications and books in which thermochemical data may be found. The following list summarizes the more important sources to which reference should be made in looking up such data.

1. International Critical Tables, Vol. V, McGraw-Hill Book Company, Inc., New York, 1929. A summary of thermochemical data available up to about the year 1927.
2. Bichowsky and Rossini, "The Thermochemistry of Chemical Substances," Reinhold Publishing Corporation, New York, 1936. A recent and critical compilation of the heats of formation of inorganic substances and the simpler carbon compounds.
3. Parks and Huffman, "The Free Energies of Some Organic Compounds," Reinhold Publishing Corporation, New York, 1932. Data on various carbon compounds.

4. Kharasch, Bureau of Standards Journal of Research, **2**, 359–430 (1929). A summary of heats of combustion of organic compounds.
5. "Annual Tables of Constants," McGraw-Hill Book Company, Inc., New York. In these tables may be found all types of recent data of thermochemical interest.
6. K. K. Kelley, U. S. Bureau of Mines Bulletins, No. 324, 371, 383, 384, 393, 406, 407. These bulletins contain valuable thermodynamic data on inorganic compounds.

Other information may be found in various chemical and chemical engineering handbooks, and in the current chemical literature.

REFERENCES FOR FURTHER READING

1. Bichowsky and Rossini, "The Thermochemistry of Chemical Substances," Reinhold Publishing Corporation, New York, 1936.
2. J. A. V. Butler, "Fundamentals of Chemical Thermodynamics," The Macmillan Company, New York, 1934–1935.
3. Hougen and Watson, "Industrial Chemical Calculations," John Wiley & Sons, Inc., New York, 1936.
4. Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.
5. J. K. Roberts, "Heat and Thermodynamics," Blackie & Son, Ltd., Glasgow, 1928.
6. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931, Vol. I.

PROBLEMS

1. The molar heat of combustion of naphthalene (M.W. = 128.16) is -1231.8 kilocal per mole. If 0.550 g of naphthalene burned in a calorimeter cause a rise in temperature of 2.050° C, what is the total heat capacity of the calorimeter?
Ans. 2544 cal/° C

2. If 1.52 g of an organic compound burned in the calorimeter in problem 1 cause the temperature to rise 1.845° C, what is the heat of combustion in calories per gram of the compound?

3. Calculate the difference between ΔH and ΔE at 25° C for the reaction
$$CH_3COOH(l) + 2 O_2(g) = 2 CO_2(g) + 2 H_2O(l) \qquad Ans. \text{ Zero}$$

4. A 0.500-g sample of heptane burned in a bomb calorimeter causes a temperature rise of 2.94° C. If the heat capacity of the calorimeter and its accessories is 1962 cal/degree and the mean temperature of the calorimeter is 25° C, calculate the heat of combustion of heptane at 25° C in calories per mole.

5. For the following reactions state whether ΔH will be significantly different from ΔE, and tell whether ΔH will be greater or less than ΔE in each case. Assume that all reactants and products are in their normal states at 25° C.

    (a) The complete combustion of sucrose ($C_{12}H_{22}O_{11}$)
    (b) The oxidation of solid naphthalene ($C_{10}H_8$) with gaseous O_2 to solid phthalic acid, $C_6H_4(COOH)_2$
    (c) The complete combustion of ethyl alcohol
    (d) The oxidation of PbS with O_2 to PbO and SO_2.

6. From the heat of combustion of *n*-butane in Table 2 calculate the heat of formation of this compound per mole at 25° C.

7. From the following equations and heats of reaction, calculate the molar heat of formation of AgCl at 25° C.

$$Ag_2O(s) + 2\,HCl(g) = 2\,AgCl(s) + H_2O(l) \quad \Delta H = -\,78,100 \text{ cal}$$

$$2\,Ag(s) + \frac{1}{2}\,O_2(g) = Ag_2O(s) \qquad\qquad \Delta H = -\,6900 \text{ cal}$$

$$\frac{1}{2}\,H_2(g) + \frac{1}{2}\,Cl_2(g) = HCl(g) \qquad\qquad \Delta H = -\,22,000 \text{ cal}$$

$$H_2(g) + \frac{1}{2}\,O_2(g) = H_2O(l) \qquad\qquad \Delta H = -\,68,400 \text{ cal}$$
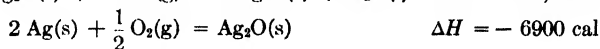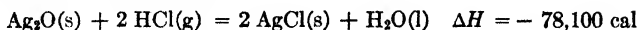
Ans. $-\,30,300$ cal

8. From the data of Table 1 calculate the heats of the following reactions at 25° C:
   (a) $Fe_2O_3(s) + CO(g) = CO_2(g) + 2\,FeO(s)$
   (b) $2\,NO_2(g) = 2\,NO(g) + O_2(g)$
   (c) $3\,C_2H_2(g) = C_6H_6(l)$
   (d) $HgO(s) + 2\,HCl(g) = H_2O(l) + HgCl_2(s)$.
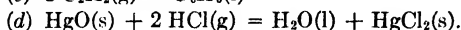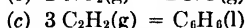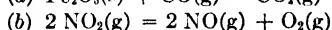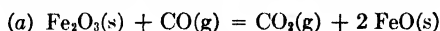
9. Using Thornton's rule, estimate the heat of combustion at constant volume of *n*-decane.

10. The following compounds originally in dilute aqueous solution are mixed together as indicated:
    (a) $CH_3COONa + HCl \longrightarrow$
    (b) $KCl + MgSO_4 \longrightarrow$
    (c) $NH_4Cl + Ba(OH)_2 \longrightarrow$
    (d) $CaCl_2 + K_2CO_3 \longrightarrow$
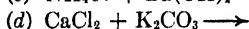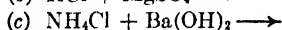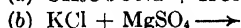
    In which cases may an appreciable thermal effect be expected? Explain your answer.

11. A 150.0-cc portion of a 0.40 N HCl is neutralized with an excess of NH_4OH in a Dewar vessel with a resulting rise in temperature of 2.36° C. If the heat capacity of the Dewar and its contents after the reaction is 315 cal/degree, calculate the heat of neutralization in calories per mole.

12. Anhydrous Na_2CO_3 dissolves in a large excess of H_2O with the evolution of 5500 cal of heat per mole. The heat of formation of $Na_2CO_3(s)$ is $-\,271,000$ cal/mole. Calculate the heat of formation per mole of Na_2CO_3 in a dilute solution.   *Ans.* $-\,276,500$ cal/mole

13. From the following reactions and thermal data

$$2\,Fe(s) + \frac{3}{2}\,O_2(g) = Fe_2O_3(s) \qquad \Delta H = -\,198,500 \text{ cal}$$

$$2\,FeO(s) + \frac{1}{2}\,O_2(g) = Fe_2O_3(s) \qquad \Delta H = -\,69,900 \text{ cal}$$

$$Fe(s) + 2\,H^+ = Fe^{++} + H_2(g) \qquad \Delta H = -\,20,800 \text{ cal}$$

$$\frac{1}{2}\,H_2(g) = H^+ \qquad\qquad\qquad \Delta H = 0 \text{ cal}$$

$$H_2(g) + \frac{1}{2}\,O_2(g) = H_2O(l) \qquad \Delta H = -\,68,400 \text{ cal}$$

calculate the heat of the reaction:

$$FeO(s) + 2\,H^+ = H_2O(l) + Fe^{++}$$
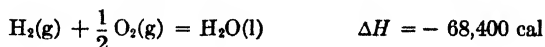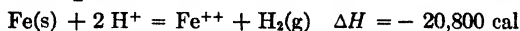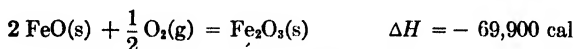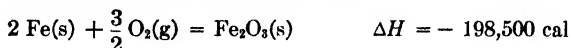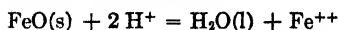
14. From the data of Tables 1 and 5, calculate the heats of the following reactions:
    (a) $Ca^{++} + CO_3^{--} = CaCO_3(s)$
    (b) $CO_3^- + 2\,H^+ = H_2O(l) + CO_2(g)$
    (c) $ZnS(s) + 2\,H^+ = Zn^{++} + H_2S(g)$
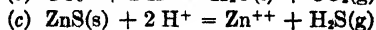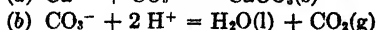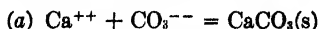
15. By interpolation from Table 6, estimate the heat effect resulting, (a) when 100 g of H_2SO_4 are added to 100 g of H_2O, and (b) when 100 g of H_2O are added to 100 g of solution containing 70% by weight of H_2SO_4.

16. From Table 6 calculate the integral heat of solution of 2 moles of H_2SO_4 in 3 moles of H_2O.

17. At 18° C the heat of solution of anhydrous $CuSO_4$ in water is $-$ 15,800 cal/mole, while that of $CuSO_4 \cdot 5\ H_2O$ is 2750 cal/mole. What is the heat of the reaction

$$CuSO_4(s) + 5\ H_2O(l) = CuSO_4 \cdot 5\ H_2O(s)$$
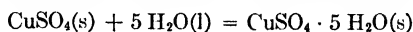
at 18° C?

18. For gaseous CO_2, $C_p = 7.0 + 0.0071\ T - 0.00000186\ T^2$ cal/mole-degree. Calculate the amount of heat necessary to raise the temperature of 200 g of gaseous CO_2 from 27° C to 227° C, (a) at constant pressure, (b) at constant volume. Assume $C_p - C_v = R$ per mole.   *Ans.* (a) 1907 cal; (b) 1509 cal

19. One mole of N_2 and 3 moles of H_2 are heated to 450° C and subjected to a pressure which results in 0.1 mole conversion of N_2 into NH_3. The gases are then cooled rapidly back to 25° C. From the following thermal data find whether heat is given up or absorbed in the whole process, and the amount of heat involved:

$$\frac{1}{2}\ N_2(g) + \frac{3}{2}\ H_2(g) = NH_3(g) \quad \Delta H_{25° C} = -\ 11,050\ cal$$
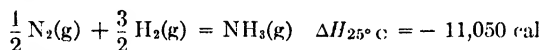
20. A mixture of gases contains 40% of CO_2, 30% of CO, and 30% of N_2 by volume. Calculate the amount of heat necessary to raise the temperature of 1000 g of this mixture from 27° C to 227° C at constant pressure.

$$N_2 : \quad C_p = 6.5 + 0.0010\ T$$
$$CO_2 : \quad C_p = 7.0 + 0.0071\ T - 0.00000186\ T^2$$
$$CO : \quad C_p = 6.6 + 0.0012\ T$$

21. Calculate the heat of vaporization of H_2O at 120° C and 1 atm pressure. The heat capacity of $H_2O(l)$ may be taken as 1.0 cal/gram-degree, C_p for the vapor as 0.45 cal/gram-degree, and the heat of vaporization at 100° C as 536 cal/gram.

22. Calculate ΔH at 1000° C for the reaction

$$2\ CO(g) + O_2(g) = 2\ CO_2(g)$$
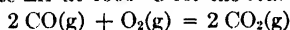
given the data:

$$\Delta H_{293° K} = -\ 135,800\ cal$$
$$CO : \quad C_p = 6.6 + 1.2 \times 10^{-3}\ T$$
$$CO_2 : \quad C_p = 7.0 + 7.1 \times 10^{-3}\ T - 1.86 \times 10^{-6}\ T^2$$
$$O_2 : \quad C_p = 6.76 + 0.606 \times 10^{-3}\ T + 0.13 \times 10^{-6}\ T^2$$
$$Ans. \quad -\ 136,240\ cal$$

23. Find an expression for ΔH as a function of T for the following reaction

$$CO_2(g) + C\ (graphite) = 2\ CO(g)$$
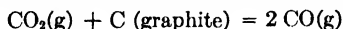
given that: $\Delta H_{293° K} = 41,400\ cal$

$$C\ (graphite) : \quad C_p = 2.672 + 2.62 \times 10^{-3}\ T - \frac{1.17 \times 10^5}{T^2}$$

See preceding problem for C_p of CO and CO_2.

Ans. $\Delta H = 40,704 + 3.53\ T - 3.66 \times 10^{-3}\ T^2 + 0.62 \times 10^{-6}\ T^3 - \dfrac{1.17 \times 10^5}{T}$

24. The expression for ΔH of formation for CO_2 as a function of temperature is

$$\Delta H = -\ 93,480 - 0.603\ T - 0.675 \times 10^{-4}\ T^2 - \frac{1.091 \times 10^5}{T}$$

Find ΔC_p for this reaction as a function of T.

25. Find an expression for ΔH as a function of T for the reaction

$$N_2(g) + O_2(g) = 2 NO(g)$$
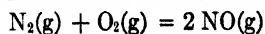
given the data:

$$\Delta H_{293° \text{ K}} = 43{,}000 \text{ cal}$$

$$\text{NO}: \quad C_p = 8.05 + 2.3 \times 10^{-4}\, T - \frac{1.56 \times 10^5}{T^2}$$

$$O_2: \quad C_p = 6.8 + 0.61 \times 10^{-3}\, T + 0.13 \times 10^{-6}\, T^2$$

$$N_2: \quad C_p = 6.5 + 1 \times 10^{-3}\, T$$

Entropy and the Third Law of Thermodynamics

The Unavailable Energy of a System. Experience with various forms of energy has established the fact that all forms of energy other than heat may be converted *completely* to heat. Thus, a paddle wheel immersed in water and agitated mechanically will generate in the water a quantity of heat exactly equivalent to the mechanical energy expended. Again, if electrical energy be supplied to a resistance, it is found that the amount of heat generated by the current passing through the resistance is exactly equal to the electrical energy consumed. The heat obtained is always equal to the work done, indicating that work may be converted completely into heat under the proper conditions.

The reverse process, the conversion of heat to work, is never complete without permanent changes in some part of the system. This, of course, is the essence of the second law of thermodynamics. Under isothermal conditions heat cannot be converted to work unless some definite and permanent compensating change in the system has occurred. For instance, an ideal gas may absorb isothermally a quantity of heat q and yield in turn an amount of work exactly equal to the heat absorbed, but this is possible only if the gas is expanded from an initial volume V_1 to a final volume V_2. Without such a compensating change no heat can be converted to work under isothermal conditions.

Under nonisothermal conditions only a part of the heat can be recovered as work, the rest being unavailable for this purpose. The amount of work obtainable from a certain quantity of heat depends on the temperature at which the heat is absorbed, the operating temperature difference, as well as on the manner of conducting the process. Under *ideal conditions of operation*, namely, *complete reversibility* in absorption of heat and performance of work, the result of the Carnot cycle gives as the maximum amount of work which may be recovered from a given quantity of heat

$$w_n = q\left(\frac{T_2 - T_1}{T_2}\right) \tag{1}$$

According to equation (1) w_m can equal q only when $T_1 = 0$. Since, however, T_1 is always greater than zero, the work recovered from a given

quantity of heat is always less than q. This indicates that even under the most ideal conditions heat cannot be converted completely to work, a part of it always remaining unavailable for doing work. It should be emphasized that this incomplete conversion of heat to work is in no way associated with any limitations or imperfections in equipment or apparatus used, but is inherent in the very nature of heat.

The total energy which any system possesses is composed in general of thermal energy due to molecular agitation plus the energy due to rotation, vibration, and electronic distribution in the molecules; i.e., some of the energy of a system is present as heat, while some of it is due to the relative positions and attractions of the atoms and molecules. Since heat cannot be converted completely into work, it must of necessity follow that the total energy of a system cannot all be recovered as work, and that some of it must always remain unavailable for this purpose. We might expect, further, that for any given state of a system there will be a definite and fixed distribution of the total energy between available and unavailable energy, and that this distribution will depend on the nature of the system and the temperature and pressure, or the temperature and volume. Such is actually the case.

Entropy and Entropy Change. The energy available for doing work which a system possesses at any given temperature is the maximum work content A, defined in Chapter II. If we define now by S *the energy per degree of absolute temperature that cannot be recovered as work*, then the amount of unavailable energy present is TS, and the total energy E is given by

$$E = A + TS \tag{2}$$

S is called the *entropy of the system*, and, in line with the above definition, represents the *unavailable energy of the system per degree* at the temperature T. Since E and A are properties of the state of the system only, it must follow that the entropy is also a property characteristic of the state of the system, and is in no way dependent on the manner in which the particular state has been achieved. In this respect the entropy has the same attributes as the functions E, H, and A; i.e., the difference in entropy between any two states is the difference between the entropy in the final state and the entropy in the initial state.

If the internal energy, work content, and entropy of a system in the initial state be designated by the subscript $_1$, while the values of the same quantities in some other state be designated by the subscript $_2$, then the change in the internal energy at constant temperature on passing from the initial state to the final state is given by

$$\Delta E = E_2 - E_1 = (A_2 + TS_2) - (A_1 + TS_1)$$
$$= (A_2 - A_1) + T(S_2 - S_1)$$

or $\qquad \Delta E = \Delta A + T\Delta S \tag{3}$

ΔS is the change in entropy between the final and initial states, and is defined by

$$\Delta S = S_2 - S_1 \tag{4}$$

Equation (3), frequently referred to as the *Gibbs-Helmholtz equation*, is very important, for it relates directly the changes in the internal energy, work content, and the entropy for any process conducted *reversibly* at the constant temperature T. Knowing any two of these quantities, the third may readily be calculated by its use.

Another fundamental equation for the entropy change of a process may be obtained by comparing equation (3) with equation (34) of Chapter II, namely,

$$\Delta E = \Delta A + q \tag{5}$$

where q is the heat absorbed reversibly in the given process. From these two equations it follows that:

$$T\Delta S = q_{\text{rev}}$$

$$\Delta S = \frac{q_{\text{rev}}}{T} \tag{6}$$

For any isothermal reversible process in which an amount of heat q is absorbed at a temperature T, therefore, the entropy change involved in the process is given simply by the heat absorbed divided by the absolute temperature at which the heat absorption takes place. When q is positive, i.e., heat is absorbed, the change in entropy is also positive, indicating an increase in unavailable energy because of the change occurring in the system. On the other hand, when q is negative, indicating an evolution of heat, ΔS is also negative, and there is a decrease in the unavailable energy of the system.

Entropies and entropy changes are expressed usually in calories per degree per given amount of substance. The quantity *calorie per degree* is called an *entropy unit* (eu).

Entropy Change in Reversible and Irreversible Isothermal Processes. Since the entropy is a characteristic property depending only on the state of the system, the entropy change suffered by a system due to some change within the system must be the same no matter how the change is accomplished. In other words, *the magnitude of the entropy change* will depend only on the initial and final states, and *will be the same whether the process is conducted reversibly or not.* Thus, in an isothermal reversible process $\Delta S = q_{\text{rev}}/T$, and this must also be the entropy change when the process is conducted irreversibly. The actual value of q in an irreversible process depends on the manner of conducting the process, and will vary anywhere from $q = 0$ to $q = q_{\text{rev}}$ when complete reversibility obtains. It is evident, therefore, that for an irreversible process

$\Delta S \neq q/T$, where q is the actual heat absorbed or evolved in the process; for, if ΔS were equal to q/T in all these cases, we should obtain a series of different values of ΔS between any two given states of the system. As this is an impossible state of affairs for a property characteristic of the states of the system only, it must follow that even in an irreversible process ΔS must be given by q_{rev}/T. To calculate, then, the entropy change involved in an irreversible process we must find first what the thermal change would have been had the process taken place reversibly, and then use equation (6). Any value of q other than q_{rev} when substituted in equation (6) will not yield ΔS.

These considerations can best be illustrated by the following two examples, the first involving an isothermal reversible process, the second an isothermal irreversible process.

(a) Calculate the entropy change involved in the isothermal reversible expansion of one mole of an ideal gas from a volume of one liter to a volume of 10 liters at 300° K. Since the process is isothermal, $\Delta E = 0$. Hence by the first law of thermodynamics $q = w$. Further, since the process is reversible, the work done must be w_m, and therefore $w_m = q_{rev}$. The entropy change for this process is given, then, according to equation (6) by

$$\Delta S = \frac{q_{rev}}{T} = \frac{w_m}{T}$$

But, according to equation (31), Chapter II, w_m is

$$w_m = nRT \ln \frac{V_2}{V_1}$$

Therefore, $$\Delta S = nR \ln \frac{V_2}{V_1} = 1.987 \times 2.303 \log \frac{10}{1}$$
$$= 4.58 \text{ cal/degree/mole}$$
$$= 4.58 \text{ entropy units (eu)/mole}$$

(b) Calculate the entropy change involved in the isothermal expansion of 1 mole of gas against a constant pressure of 1 atmosphere from a volume of 1 liter to 10 liters at 300° K. The expansion described in this problem is irreversible, but again isothermal. For such a process q again equals w, but w now is not the maximum work, and hence q does not equal q_{rev}. To find q_{rev} for substitution in equation (6) we must imagine the same process conducted reversibly. For the reversible process, however, q_{rev} is given by the same equation as in the preceding problem, and consequently we have again that

$$\Delta S = nR \ln \frac{V_2}{V_1}$$
$$= 2.303 \, R \log \frac{10}{1}$$
$$= 4.58 \text{ eu/mole}$$

The entropy change is the same for both processes because the initial and final states are the same, although different paths were followed in the two cases in passing from one point to the other.

Entropy Change in Reversible and Irreversible Nonisothermal Processes. Any nonisothermal process may be considered to be composed of a successive series of isothermal processes, each occurring at a temperature infinitesimally different from the preceding. If each of these isothermal processes takes place reversibly, the entropy change for each will be given by the heat absorbed in each of these steps, q_{rev}, divided by T, the temperature at which the heat absorption takes place. The total entropy change for a process occurring between the temperatures T_1 and T_2 will be, then, the sum of the small continuous isothermal entropy changes, or

$$\Delta S = \sum_{T_1}^{T_2} \frac{q_{rev}}{T} \tag{7}$$

From what has been said before it is evident that equation (7) is applicable also to irreversible nonisothermal processes provided the q's employed are not the heats actually observed, but the heats evaluated for the corresponding reversible processes between the same two states.

Entropy Change and Irreversibility. It can readily be shown that, when a system is subjected to a series of reversible operations such that at the end the system is returned to its initial conditions, the change in entropy for the complete cycle is zero. As an instance in point consider the Carnot cycle. The total change in entropy for the whole cycle must be the sum of the entropy changes involved in the four stages of the cycle. In the first stage the system absorbs isothermally and reversibly a quantity of heat q_2 at T_2, and consequently the entropy change is $\Delta S_1 = q_2/T_2$. Similarly the entropy change for the third stage, which involves the isothermal and reversible evolution of q_1 calories of heat at T_1, is $\Delta S_3 = -q_1/T_1$. Since the second and fourth stages of the cycle are adiabatic, no heat is evolved or absorbed, and hence for these stages $\Delta S_2 = \Delta S_4 = 0$. The total change in entropy for the complete Carnot cycle is, then,

$$\Delta S = \frac{q_2}{T_2} - \frac{q_1}{T_1} \tag{8}$$

However, the maximum work performed in the cycle is equal to the difference between the heat absorbed at T_2 and the heat evolved at T_1, or $w_m = q_2 - q_1$. Therefore,

$$w_m = q_2 - q_1 = q_2 \frac{(T_2 - T_1)}{T_2}$$

$$q_2 - q_1 = q_2 - \frac{q_2 T_1}{T_2}$$

$$\frac{q_2}{T_2} - \frac{q_1}{T_1} = 0 \tag{9}$$

Inserting equation (9) in equation (8), we see that ΔS for the whole cycle is

$$\Delta S = 0 \tag{10}$$

This result is valid for any cyclical process, provided the whole process in all its phases is conducted reversibly. If either the whole cycle or any part of it involves any irreversibility, however, it can be shown that the change in entropy for the cycle is no longer zero for the system and its surroundings, but some positive value whose magnitude depends on the degree of the irreversibility. The entropy change for a cycle constitutes, then, a good criterion of the reversibility of the processes involved in the cycle. When $\Delta S = 0$, the cycle is completely reversible; when, on the other hand, ΔS for the cycle is positive, some irreversibility exists, and its degree may be estimated from the magnitude of the entropy change.

The above deductions regarding entropy changes in reversible and irreversible cycles apply also to all other types of processes, whether cyclical or not, provided the entropy changes in the operating body as well as in its surroundings are considered. In any system composed of a body and its surroundings, all completely reversible changes will occur without any modification in total entropy. To illustrate the validity of this statement, consider the isothermal reversible expansion of an ideal gas at temperature T. We have seen that for such an expansion $\Delta S_{gas} = q_{rev}/T$, where q_{rev} is the heat absorbed reversibly by the gas at T from its surroundings, also at T. But, the surroundings, by losing a quantity of heat q_{rev} at T, experience an entropy decrease equal to $- q_{rev}/T$. Consequently the entropy increase of the gas is compensated completely by the entropy decrease of the surroundings, and the entropy change for the system as a whole is zero.

On the other hand, any process occurring irreversibly in a system composed of a body and its surroundings must be attended by an increase in total entropy. To show this, consider the same example, but let the expansion be irreversible. Although in expanding to the same final volume the gas will absorb this time only a quantity of heat q which is less than q_{rev}, its entropy change will be the same as before, namely, q_{rev}/T. However, the entropy decrease suffered by the surroundings will no longer be $- q_{rev}/T$, but $- q/T$. Since q/T is necessarily less than q_{rev}/T, it must follow that the entropy change for the entire process, $\Delta S = q_{rev}/T - q/T$, must be greater than zero. In other words, the irreversible process must be attended by an increase in total entropy. Furthermore, since the difference between these two entropy changes will increase with the extent of irreversibility, the total ΔS will also increase, and it will be greater the greater the degree of irreversibility.

The Second Law of Thermodynamics and Entropy. Completely reversible processes, involving as they do a balance of driving and opposing forces, must of necessity occur very slowly. In fact, to carry out any change in a completely reversible manner would require infinite time. Consequently, any process that does take place in a finite time must be irreversible, and it must be attended by an increase in the total entropy of all the bodies involved. The latter conclusion permits a statement of the second law of thermodynamics in its most general form, namely, *that all processes in nature tend to occur only with an increase in entropy, and that the direction of change is always such as to lead to the entropy increase.* The several forms of the second law given in Chapter II are deducible from the above, and are merely special limited cases of this general statement.

Variation of Entropy with Temperature. A system with an entropy S at a temperature T will in general experience both a change in temperature and a change in entropy on absorption of heat. Due to the absorption of the infinitesimally small quantity of heat, q_{rev}, at the temperature T the system will experience an entropy change dS such that

$$dS = \frac{q_{rev}}{T} \tag{11}$$

When the heat is absorbed at constant pressure, q_{rev} is equal to the heat capacity of the system C_p, multiplied by the temperature change dT, or $q_{rev} = C_p\, dT$. Inserting this value of q_{rev} in equation (11) we obtain

$$(dS)_P = \frac{q_{rev}}{T} = \frac{C_p\, dT}{T}$$

and hence,
$$\left(\frac{\partial S}{\partial T}\right)_P = \frac{C_p}{T} \tag{12}$$

Similarly, if the heat be absorbed at constant volume, $q_{rev} = C_v\, dT$, and equation (11) becomes

$$(dS)_V = \frac{q_{rev}}{T} = \frac{C_v\, dT}{T}$$

or
$$\left(\frac{\partial S}{\partial T}\right)_V = \frac{C_v}{T} \tag{13}$$

Equations (12) and (13) are very significant, for they give the variation of the entropy with temperature at constant pressure and constant volume, and permit the evaluation of the entropy change between any two temperatures. To obtain the entropy change on passing from a temperature T_1 to T_2 we need only integrate these expressions. From equa-

tion (12) the entropy change at constant pressure between T_1 and T_2 follows as

$$\left(\frac{\partial S}{\partial T}\right)_P = \frac{C_p}{T}$$

$$dS = \frac{C_p\, dT}{T}$$

$$\Delta S = S_2 - S_1 = \int_{T_1}^{T_2} \frac{C_p\, dT}{T} \tag{14}$$

When C_p can be considered constant over the given temperature interval equation (14) reduces to

$$\Delta S = \int_{T_1}^{T_2} \frac{C_p\, dT}{T} = C_p \int_{T_1}^{T_2} d \ln T$$

$$= C_p \ln \frac{T_2}{T_1} \tag{15}$$

When this approximation is not valid, however, C_p as a function of T must be substituted in equation (14), and the integration must be carried out term by term between the given limits. In the same manner equation (13) leads to the integral,

$$\Delta S = S_2 - S_1 = \int_{T_1}^{T_2} \frac{C_v\, dT}{T} \tag{16}$$

which at constant C_v reduces to

$$\Delta S = C_v \ln \frac{T_2}{T_1} \tag{17}$$

To illustrate the procedure involved, consider the problem of finding the increase in entropy suffered by 2 moles of hydrogen on being heated at constant pressure from 300° K to 600° K. The heat capacity *per mole* of the gas is given by the equation

$$H_2(g): \quad C_p = 6.50 + 0.0009\, T \text{ cal/mole}$$

Applying equation (14), we have for ΔS

$$\Delta S = S_{600°\,\mathrm{K}} - S_{300°\,\mathrm{K}} = \int_{300°\,\mathrm{K}}^{600°\,\mathrm{K}} \frac{C_p\, dT}{T}$$

$$= \int_{300°\,\mathrm{K}}^{600°\,\mathrm{K}} \frac{2(6.50 + 0.009\, T)dT}{T}$$

$$= 13.0 \times 2.30 \log \frac{600}{300} + 0.0018(600 - 300)$$

$$= 10.47 \text{ eu}$$

The arguments developed so far and the equations deduced are generally applicable to any kind of system and to any type of change occurring within a system. Now the discussion will be extended to some specific instances, namely, to the entropy changes taking place both in ideal gases and in physical and chemical transformations.

Change of Entropy in Perfect Gases. The entropy of any gas, being a function of the state of the system only, is in general determined by any two of the three variables P, T, and V. Whenever any of these variables are changed, the entropy of the gas is also changed. The manner in which the total entropy of an ideal gas changes with these variables may readily be determined from thermodynamics. The variables chosen for consideration will be (1) temperature and volume, and (2) temperature and pressure, as these are the ones commonly encountered in calculations.

(1) Variable Temperature and Volume. According to the first law of thermodynamics, when a system absorbs an infinitesimally small quantity of heat q, this heat is utilized to increase the internal energy of the system by the amount dE, and to do the work w, or

$$q = dE + w \tag{18}$$

If the absorption of heat and performance of work are carried out reversibly, then, according to equation (11), $q = TdS$, and equation (18) becomes

$$TdS = dE + w \qquad \cdot$$

and
$$dS = \frac{dE + w}{T} \tag{19}$$

For a system containing n moles of gas with a heat capacity at constant volume of C_v per mole, $dE = nC_v\, dT$. Again, the work done by a gas is $w = PdV$, and since for an ideal gas $P = nRT/V$, then $w = nR\, TdV/V$. Inserting these values for dE and w in equation (19), dS becomes

$$dS = \frac{nC_v\, dT}{T} + \frac{nRdV}{V} \tag{20}$$

On integrating this equation between the limits T_1 and T_2 and V_1 and V_2, considering C_v constant, we obtain, finally,

$$\Delta S = S_2 - S_1 = nC_v \ln \frac{T_2}{T_1} + nR \ln \frac{V_2}{V_1} \tag{21}$$

Equation (21) gives the total change in entropy attendant upon changing the temperature of an ideal gas from T_1 to T_2 and the volume from V_1 and V_2. The entropy change is composed of two terms, one due to the change of temperature only, the other to change of volume only.

When there is no change in temperature, the first term on the right of equation (21) is zero, and ΔS is given by

$$\Delta S = nR \ln \frac{V_2}{V_1} \tag{22}$$

This equation gives the change in entropy which takes place when only the volume of the ideal gas is changed from V_1 to V_2 *at constant temperature.* On the other hand, when the volume of the system remains constant and only the temperature is varied, the second term on the right in equation (21) is zero, and the equation reduces to equation (17).

(2) *Variable Temperature and Pressure.* If the equalities $dE = nC_v \, dT$ and $w = PdV$ are substituted in equation (19), then

$$dS = \frac{nC_v \, dT + PdV}{T} \tag{23}$$

For an ideal gas $PV = nRT$. Differentiating this expression totally, we get

$$PdV + VdP = nRdT$$
$$PdV = nRdT - VdP$$

and hence,

$$dS = \frac{nC_v \, dT + nRdT - VdP}{T}$$

$$= \frac{n(C_v + R)dT}{T} - \frac{VdP}{T} \tag{24}$$

However, $C_v + R = C_p$, and $V = nRT/P$. Therefore,

$$dS = \frac{nC_p \, dT}{T} - \frac{nRdP}{P} \tag{25}$$

On integration of equation (25) at constant C_p between the limits T_1 and T_2 and P_1 and P_2, we obtain

$$\Delta S = S_2 - S_1 = nC_p \ln \frac{T_2}{T_1} - nR \ln \frac{P_2}{P_1} \tag{26}$$

This equation gives the total change in entropy occasioned on changing both the temperature and pressure of the ideal gas. The first term on the right gives the change in entropy due to change in temperature only, and is identical with equation (15). The second term, on the other hand, gives the change in entropy due to change of pressure at *constant temperature,* and hence this change is given by

$$\Delta S = -nR \ln \frac{P_2}{P_1} = nR \ln \frac{P_1}{P_2} \tag{27}$$

Equations (21) and (26) constitute the essential equations required for calculating the entropy changes suffered by ideal gases under most conditions encountered. Thus, when ΔS due to change in temperature at constant volume is desired, the second term in equation (21) drops out, and the equation reduces to $\Delta S = nC_v \ln T_2/T_1$. If, on the other hand, ΔS at constant temperature due to change in volume is sought, the first term is zero, and the equation reduces to equation (22). When both T and V vary, the full equation must be used. The same considerations apply to equation (26), and the variation of entropy with temperature at constant pressure, with pressure at constant temperature, and with both temperature and pressure, as may be seen from the following example.

Suppose it is desired to find ΔS for 5 moles of an ideal, monatomic gas which is being compressed by heating from a pressure of 1 atmosphere at 0° C to 10 atmospheres at 25° C. In this instance both the pressure and temperature are varied, and hence equation (26) must be used in full. Since for an ideal monatomic gas $C_p = 5/2\,R$, we have

$$\Delta S = nC_p \ln \frac{T_2}{T_1} - nR \ln \frac{P_2}{P_1}$$

$$\Delta S = 5 \times \frac{5}{2}\,(1.987) \times 2.303 \log \frac{298}{273} - 5 \times 1.987 \times 2.303 \log \frac{10}{1}$$

$$= 2.17 - 22.86$$

$$= -\,20.69 \text{ cal/degree (or eu)/5 moles}$$

It is readily seen that ΔS due to the change in temperature is 2.17 eu, while ΔS due to change in pressure is $-\,22.85$ eu. The total entropy change must of course be the algebraic sum of the two, or $-\,20.69$ eu.

Entropy Change in Physical Transformations. Changes in entropy accompany not only variations in the temperature, pressure, or volume of a system, but also physical transformations such as fusion, vaporization, or transition from one crystalline form to another. All these processes may be represented by the general scheme

$$B_1 = B_2$$

and, for all, the change in entropy is defined as

$$\Delta S = S_2 - S_1 \tag{28}$$

where S_2 is the entropy of the final form and S_1 the entropy of the initial form.

It has been shown in preceding chapters that all the transitions enumerated take place with a volume change from V_1 to V_2 at constant temperature T and pressure P, and are accompanied by an exchange of q

calories of heat between the system and its surroundings. The work involved in such processes is obviously

$$w = -\Delta A = P(V_2 - V_1)$$

and hence, according to the first law,

$$\begin{aligned}
q = \Delta E + w &= \Delta E - \Delta A \\
&= \Delta E + P(V_2 - V_1) \\
&= \Delta H
\end{aligned}$$

But in view of equation (3), $\Delta E - \Delta A = T\Delta S$. Consequently,

$$T\Delta S = q = \Delta H$$
$$\Delta S = \frac{q}{T} = \frac{\Delta H}{T} \tag{29}$$

For all transformations of state, therefore, the change in entropy is given by the heat necessary to accomplish the transition divided by the temperature at which the transition takes place. Of necessity equation (29) is valid only when reversible conditions obtain during the transformation, i.e., when equilibrium exists between the two forms.

As an illustration of the use of equation (29), consider the problem of finding the entropy difference for the transition

$$H_2O(l, 1 \text{ atm}) = H_2O(g, 1 \text{ atm}) \qquad \Delta H_{373.2°K} = 9710 \text{ cal/mole}$$
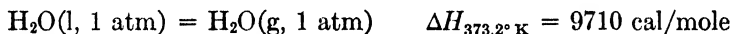
Since, at 373.2° K and 1 atmosphere pressure, the normal boiling point of water, $H_2O(l)$ is in equilibrium with $H_2O(g)$, then

$$\begin{aligned}
\Delta S = S_g - S_l = \frac{\Delta H}{T} \\
= \frac{9710}{373.2} \\
= 26.03 \text{ cal/degree/mole}
\end{aligned}$$

Entropy changes may also be calculated for nonequilibrium and even irreversible transitions, but the change in entropy will no longer be given by equation (29). Consider first the reversible but nonequilibrium process at 373.2° K:

$$H_2O(l, 1 \text{ atm}) = H_2O(g, 0.1 \text{ atm})$$
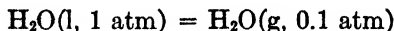
Since the change in entropy does not depend on the manner of accomplishing the change, we may first vaporize the water isothermally and reversibly to steam at 1 atmosphere pressure, and then expand the steam

reversibly and isothermally from 1 to 0.1 atmosphere. The total change in entropy will then be

$$\Delta S = \Delta S_{\text{vaporization}} + \Delta S_{\text{expansion}}$$
$$= \frac{\Delta H}{T} + R \ln \frac{P_1}{P_2}$$
$$= \frac{9710}{373.2} + 4.58 \log \frac{1}{0.1}$$
$$= 30.61 \text{ eu/mole}$$

Again, suppose it is desired to calculate the entropy change involved in an irreversible process such as the conversion of a mole of water to a mole of ice at $-15°$ C. Water and ice are in equilibrium only at $0°$ C, and, hence, equation (29) will not give the entropy change at $-15°$ C. However, if we determine the entropy changes involved in the following steps:

(1) The heating of a mole of water from $-15°$ C to $0°$ C.
(2) Reversible conversion of a mole of water to a mole of ice at $0°$ C.
(3) Cooling of a mole of ice from $0°$ C to $-15°$ C.

then, the sum of these entropy changes should be the same as the entropy change of the irreversible process $H_2O(l) = H_2O(s)$ at $-15°$ C, since the net process is the same in both cases.

Entropy Change in Chemical Reactions. The entropy change accompanying a chemical reaction is defined as the difference between the sum of the entropies of all the products and the sum of the entropies of all reactants. For any reaction between, say, b moles of B and c moles of C to yield d moles of D and e moles of E, namely,

$$bB + cC = dD + eE$$
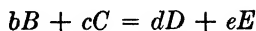
the entropy change is given by

$$\Delta S = (dS_D + eS_E) - (bS_B + cS_C) \tag{30}$$

where the symbols S_D, S_E, S_B, and S_C represent the entropy *per mole* of the species D, E, B, and C respectively.

Entropy changes of chemical reactions are evaluated at constant temperature. For any given reaction and at any given temperature the entropy change is definite and just as characteristic of the reaction as the change in the internal energy or heat content. The manner in which the entropy change of a reaction depends on the temperature at constant pressure may readily be deduced from equations (30) and (12). If equa-

tion (30) is differentiated with respect to temperature at constant pressure, then

$$\left[\frac{\partial(\Delta S)}{\partial T}\right]_P = \left[d\left(\frac{\partial S_D}{\partial T}\right)_P + e\left(\frac{\partial S_E}{\partial T}\right)_P\right] - \left[b\left(\frac{\partial S_B}{\partial T}\right)_P + c\left(\frac{\partial S_C}{\partial T}\right)_P\right]$$

But, according to equation (12) $(\partial S/\partial T)_P = C_p/T$. Hence,

$$\begin{aligned}
\left(\frac{\partial(\Delta S)}{\partial T}\right)_P &= \left(\frac{dC_{p_D}}{T} + \frac{eC_{p_E}}{T}\right) - \left(\frac{bC_{p_B}}{T} + \frac{cC_{p_C}}{T}\right) \\
&= \frac{(dC_{p_D} + eC_{p_E}) - (bC_{p_B} + cC_{p_C})}{T} \\
&= \frac{\Delta C_p}{T}
\end{aligned} \tag{31}$$

ΔC_p is the difference between the C_p's of products and reactants. On integration of equation (31) between the limits T_1 and T_2, we obtain

$$\left[\frac{\partial(\Delta S)}{\partial T}\right]_P = \frac{\Delta C_p}{T}$$

$$\int_{\Delta S_1}^{\Delta S_2} d(\Delta S) = \int_{T_1}^{T_2} \frac{\Delta C_p}{T} dT$$

$$\Delta S_2 - \Delta S_1 = \int_{T_1}^{T_2} \frac{\Delta C_p}{T} dT \tag{32}$$

ΔS_2 is the entropy change at T_2, while ΔS_1 is the entropy change of the reaction at T_1. The difference between the two is given by the integral in equation (32). If ΔC_p may be assumed constant over the temperature interval in question, then equation (32) becomes

$$\begin{aligned}
\Delta S_2 - \Delta S_1 &= \Delta C_p \int_{T_1}^{T_2} \frac{dT}{T} \\
&= \Delta C_p \ln \frac{T_2}{T_1}
\end{aligned} \tag{33}$$

If ΔC_p is not constant, however, the expression for ΔC_p as a function of T must be obtained in the same manner as described in the chapter on thermochemistry, and the integration must be carried out term by term between the given temperature limits.

The isothermal entropy change of a reaction is usually calculated either by means of equation (3) or one of its modifications, or through equation (30). To use equation (3) for ΔS, both ΔE and ΔA of the reaction at a temperature T must be known. More commonly ΔS of a reaction is calculated from ΔH and ΔF, the free energy change of the reaction, a quantity related to the entropy and described in the next chapter. On

the other hand, to obtain ΔS of a reaction with the aid of equation (30), the entropies of the individual substances at the temperature in question must be available. These are obtained at present through the *third law of thermodynamics*.

The Third Law of Thermodynamics. As a result of the researches of T. W. Richards, Walter Nernst, Max Planck, and others, another fundamental principle of thermodynamics, which deals with the entropy of pure crystalline substances at the absolute zero of temperature, has come into being. This principle, generally referred to as the third law of thermodynamics, states that the *entropy of all pure crystalline solids may be taken as zero at the absolute zero of temperature*. The statement is confined to pure crystalline solids because theoretical argument and some experimental evidence point to the fact that the entropy of solutions and supercooled liquids may not be taken as zero at $0°$ K. For pure crystalline solids the law has been verified in many instances, and at present little doubt remains as to general validity of the above statement of the law.

The importance of the third law lies in the fact that it permits the calculation of absolute values of the entropy of pure substances from heat capacity data alone. We have seen that entropy changes with temperature at constant pressure may be calculated through equation (14). Since, however, according to the third law $S = 0$ for a pure crystalline substance at $T = 0$, equation (14) may be integrated from this lower limit to any temperature T such that

$$\int_{S=0}^{S=S} dS = \int_{T=0}^{T=T} \cdot \frac{C_p \, dT}{T}$$

$$S_T = \int_0^T C_p \, d \ln T$$

$$= 2.303 \int_0^T C_p \, d \log_{10} T \tag{34}$$

S_T, known as the *absolute entropy* of the substance, is always a positive quantity. All that is necessary for the integration of equation (34) in order to obtain absolute entropies is a knowledge of the heat capacities of the substance from $T = 0$ to any temperature desired.

Calculation of Absolute Entropies. The integral in equation (34) is practically always evaluated graphically by plotting experimental C_p data vs. $\log_{10} T$ and determining the area under the curve. Since the area under the curve is equal to the integral, the entropy of any substance is given by

$$S_T = 2.303 \int_0^T C_p \, d \log_{10} T = 2.303 \left[\text{Area} \right]_{T=0}^{T=T} \tag{35}$$

In practice heat capacities are usually measured from approximately 20° K to room temperature, and extrapolation is resorted to from ca. 20° K to $T = 0$. Such extrapolations may be made graphically, but more usually are made with the aid of the Debye third power law (see
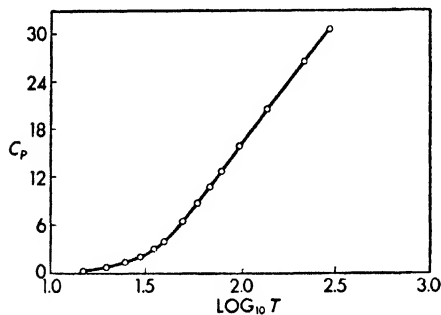
Fig. 1. Heat Capacity of Na_2SO_4 at Various Temperatures

Chapter IV), which yields heat capacities of solids at low temperatures. The uncertainty involved in such extrapolations is usually quite small, for the area between $T = 0$ and $T = $ ca. 20° is very small compared to the total area under the curve.

The data for sodium sulfate given in Table 1, and due to Pitzer and Coulter,[1] are typical of heat capacities of solids, and indicate the trend of C_p with temperature. A plot of C_p vs. $\log_{10} T$ constructed from these data is shown in Fig. 1. The area under the curve between $T = 14$ and $T = 298.15°$ K is 15.488, while the area from $T = 0$ to $T = 14$ is 0.026. Hence the absolute entropy per mole of sodium sulfate at 298.15° K is

$$Na_2SO_4(s): S_{298.15° K} = 2.303(15.488 + 0.026)$$
$$= 35.73 \text{ cal/degree/mole}$$

TABLE 1

HEAT CAPACITY OF Na_2SO_4 AT VARIOUS TEMPERATURES

$T°$ K	C_p (cal/mole/degree)	$T°$ K	C_p (cal/mole/degree)
15	0.222	60	8.743
20	0.581	70	10.85
25	1.213	80	12.76
30	1.990	100	15.93
35	2.980	140	20.50
40	4.077	220	26.51
50	6.423	300	30.51

Absolute entropies, not only of solids but also of substances that are liquid or gaseous at room temperature, can be obtained with the third law. The total absolute entropy of a substance in a particular state at a given temperature will be the sum of all the entropy changes the substance has to undergo in order to reach the particular state from the crystalline solid at absolute zero. Thus, if a substance is gaseous at

[1] Pitzer and Coulter, J. Am. Chem. Soc., **60**, 1310 (1938).

1 atmosphere pressure at 298° K, the entropy of the gas must be the sum of the entropies involved in heating the crystalline solid from $T = 0$ to $T = T_f$, the fusion point; (2) the entropy of fusion, $\Delta H_f/T_f$; (3) the entropy of heating the liquid from T_f to T_b, the normal boiling point; (4) the entropy of vaporization, $\Delta H_v/T_b$; and (5) the entropy of heating the gas from T_b to 298° K. These changes and the attendant entropies are shown in the following scheme:

$$\Delta S_1 = \int_0^{T_f} C_{p_s} d\ln T \qquad \Delta S_f = \frac{\Delta H_f}{T_f} \qquad \Delta S_2 = \int_{T_f}^{T_b} C_{p_l} d\ln T$$

Solid $\xrightarrow{\hspace{2cm}}$ Solid $\xrightarrow{\hspace{2cm}}$ Liquid $\xrightarrow{\hspace{2cm}}$

$T = 0$ $\qquad\qquad T = T_f$ $\qquad\qquad T = T_f$

$$\Delta S_v = \frac{\Delta H_v}{T_b} \qquad \Delta S_3 = \int_{T_b}^{298} C_{p_g} d\ln T$$

Liquid $\xrightarrow{\hspace{2cm}}$ Gas $\xrightarrow{\hspace{2cm}}$ Gas

$T = T_b$ $\qquad\qquad T = T_b$ $\qquad\qquad T = 298° K$

It follows, therefore, that the absolute entropy of any substance at a temperature T may be written as

$$S_T = \int_0^{T_f} C_{p_s} d\ln T + \frac{\Delta H_f}{T_f} + \int_{T_f}^{T_b} C_{p_l} d\ln T + \frac{\Delta H_v}{T_b} + \int_{T_b}^{T} C_{p_g} d\ln T \quad (36)$$

If a substance is solid at temperature T, only the first integral applies with $T_f = T$; when it is liquid, the first three terms are used with $T_b = T$ now; when it is gaseous, the full equation must be used. Equation (36) may further be complicated if the solid undergoes any crystalline modifications between $T = 0$ and $T = T_f$. When such transitions occur, the entropy or entropies of transition must be added to the others in equation (36).

The steps involved in the calculation of the absolute entropy of a substance that is gaseous at 25° C may be illustrated with the following example. Messerly and Aston [1] have shown that at 1 atmosphere pressure methyl chloride is solid between 0° K and 175.44° K, liquid between 175.44° K and 248.94° K, and gaseous thereafter, and that the entropies accompanying the heating of the solid, liquid, and gas between the temperatures indicated are

$$\Delta S_{\text{solid}} = \int_0^{175.44°} C_{p_s} d\ln T = 18.48 \text{ eu/mole}$$

$$\Delta S_{\text{liquid}} = \int_{175.44°}^{248.94°} C_{p_l} d\ln T = 6.24 \text{ eu/mole}$$

$$\Delta S_{\text{vapor}} = \int_{248.94°}^{298.15°} C_{p_g} d\ln T = 1.67 \text{ eu/mole}$$

Further, since the heat of fusion at 175.44° K is 1537 calories per mole, and the heat of vaporization at 248.94° K is 5147 calories per mole, the

[1] Messerly and Aston, J. Am. Chem. Soc., **62**, 889 (1940).

entropy changes accompanying fusion and vaporization must be

$$\Delta S_{\text{fusion}} = \frac{\Delta H_f}{T_f} = \frac{1537}{175.44} = 8.76 \text{ eu/mole}$$

$$\Delta S_{\text{vaporization}} = \frac{\Delta H_v}{T_b} = \frac{5147}{248.94} = 20.68 \text{ eu/mole}$$

Substituting these entropies in equation (36) and adding, we obtain for the absolute entropy of methyl chloride at 298.15° K

$$S_{298\ 15°\ K} = 18.48 + 8.76 + 6.24 + 20.68 + 1.67$$
$$= 55.83 \text{ eu/mole}$$

Use of Absolute Entropies in Calculations. Absolute entropies are generally calculated at 298° K. Values of these thermodynamic constants are available at present for most elements and for many compounds, both inorganic and organic. Fairly complete tabulations may be found in books on thermodynamics, in some of the references listed at the end of the chapter, and in the chemical literature.

In Table 2 are given absolute entropies at 25° C for a representative number of elements and compounds. From such individual entropies

TABLE 2

ABSOLUTE ENTROPIES OF ELEMENTS AND COMPOUNDS AT 25° C [1]

(In entropy units per gram atom or mole)

Substance	S	Substance	S
$H_2(g)$	31.23	$H_2O(l)$	16.75
C (diamond)	0.585	$H_2O(g)$	45.13
C (graphite)	1.36	$CO(g)$	47.32
$N_2(g)$	45.79	$CO_2(g)$	51.08
$O_2(g)$	49.03	$HgCl_2(s)$	34.6
Na(s)	12.2	HgCl(s)	23.5
Mg(s)	7.77	CuI(s)	23.1
S (rhombic)	7.62	AgCl(s)	23.0
S (monoclinic)	7.78	AgI(s)	27.6
$Cl_2(g)$	.53.31	$Fe_2O_3(s)$	21.5
Fe(s)	6.47	MgO(s)	6.55
Cu(s)	7.97	NaCl(s)	17.3
$Br_2(g)$	58.63	$C_2H_6(g)$	53.51
Ag(s)	10.20	$CH_3OH(l)$	30.3
$I_2(s)$	27.9	$C_2H_5OH(l)$	38.4
$I_2(g)$	62.29	$C_6H_6(l)$	41.9
Hg(l)	18.5	$C_6H_5OH(l)$	34.0
Hg(g)	41.80	$CH_3COOH(l)$	38.0

[1] From K. K. Kelley, Bureau of Mines Bulletin No. 434 (1941); Parks and Huffman, "Free Energies of Some Organic Compounds," Reinhold Publishing Corporation, New York, 1932.

the entropy changes involved in chemical reactions may readily be calculated with the aid of equation (30). Suppose the entropy change is required attending the reaction

$$C(s, \text{graphite}) + 2\,H_2(g) + \frac{1}{2}\,O_2(g) = CHO_3H(l) \quad \Delta S_{298° \text{K}} = ? \quad (37)$$
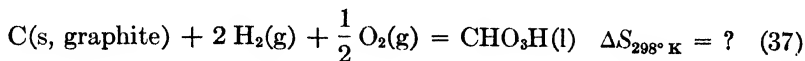
From Table 2 the entropies per mole of the various species are

$$S_{CH_3OH(l)} = 30.3 \qquad S_{H_2} = 31.23 \qquad S_{O_2} = 49.03 \qquad S_C = 1.36$$

Hence, according to equation (30),

$$\Delta S_{298° \text{K}} = S_{CH_3OH} - (S_O + 2\,S_{H_2} + \frac{1}{2}\,S_{O_2})$$
$$= 30.3 - 1.36 - 62.46 - 24.52$$
$$= -58.0 \text{ cu}$$

In a similar manner may be calculated ΔS values for other reactions at 298° K, provided the necessary absolute entropies are known. To obtain the entropy changes at temperatures other than 298° K, we need only apply equation (32). In terms of this equation the entropy change for any reaction at a temperature T in relation to ΔS at 298° is given by

$$\Delta S_T - \Delta S_{298°} = \int_{298}^{T} \frac{\Delta C_p\,dT}{T}$$
$$\Delta S_T = \Delta S_{298} + \int_{298}^{T} \frac{\Delta C_p\,dT}{T} \quad (38)$$

The evaluation of the integral has already been discussed.

REFERENCES FOR FURTHER READING

1. J. A. V. Butler, "Fundamentals of Chemical Thermodynamics," The Macmillan Company, New York, 1934–1935, 2 vols.
2. C. N. Hinshelwood, "Thermodynamics for Students of Chemistry," E. P. Dutton & Company, Inc., New York, 1926.
3. K. K. Kelley, Bureau of Mines Bulletins No. 350, 384, 394, 406, 434.
4. Kiefer and Stuart, "Principles of Engineering Thermodynamics," John Wiley & Sons, Inc., New York, 1930.
5. Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.
6. F. H. MacDougall, "Introduction to Thermodynamics and Chemistry," third edition, John Wiley & Sons, Inc., New York, 1939.
7. D. J. Martin, "Thermodynamics for Chemists," Longmans Green and Company, New York, 1933.
8. M. Planck, "Treatise on Thermodynamics," Longmans Green and Company, New York, 1927.
9. Taylor and Glasstone, "Treatise on Physical Chemistry," third edition, D. Van Nostrand Company, Inc., New York, 1942, Vol. I.

PROBLEMS

1. What is the entropy change in the isothermal compression of 200 g of N_2 from a pressure of 1 to 5 atm at 25° C? Assume that N_2 is a perfect gas.

Ans. − 22.85 eu

2. One mole of a perfect gas contained in a 10-liter vessel at 27° C is permitted to expand freely into an evacuated vessel of 10-liter capacity so that the final volume is 20 liters. What amount of work is done and what amount of heat is absorbed in the process? What is the change in entropy accompanying the process?

Ans. $\Delta S = 1.38$ eu

3. Assuming C_p for N_2 to be constant and equal to $7/2\, R$, what will be the change in entropy upon heating 10 g of the gas from 0 to 100° C, (a) at constant pressure? (b) at constant volume? $C_p - C_v = R$ per mole.

4. Given that C_p for CO_2 is $C_p = 7.70 + 5.30 \times 10^{-3}\, T - 0.83 \times 10^{-6}\, T^2$ calculate the entropy change which results upon heating 1 mole of CO_2 from 300 to 600° K at constant volume.

5. Using the data of the preceding problem, and assuming CO_2 to be an ideal gas, calculate the difference in entropy content between 0.5 mole of CO_2 at standard conditions and 0.5 mole of CO_2 at 100° C under a pressure of 3 atm.

Ans. 0.36 eu greater at 100° C

6. What is the difference in entropy between 1 mole of N_2 at standard conditions and 1 mole of N_2 at 200° C when the molar volume is 50 liters? Assume that C_p is $7/2\, R$ and that N_2 behaves ideally.

7. The atomic heat capacity of solid Mo is given by the equation

$$C_p = 5.69 + 1.88 \times 10^{-3}\, T - \frac{0.503 \times 10^5}{T^2}$$

Find the change in entropy which accompanies the heating of 1 atomic weight of Mo from 0° C to its melting point, 2620° C.

8. CH_3COOH melts at 16.6° C with a heat of fusion of 44.0 cal/gram. Calculate the entropy of fusion per mole. *Ans.* 9.13 eu

9. The heat of sublimation of CO_2 is 6220 cal/mole, while the sublimation temperature is 194.6° K. Find the entropy of sublimation of 2 moles of CO_2.

10. In the transformation $AgI(\alpha) = AgI(\beta)$ the heat of transition is + 1530 cal/mole, while the transition temperature is 146.5° C. What will be the entropy of transition?

11. One g of ice at 0° C is added to 10 g of H_2O at the boiling point. What will be the final temperature, and what is the entropy change accompanying the process? Assume that the heat of fusion of H_2O is 80 cal/gram and the specific heat 1 cal/gram.

12. What is the entropy change involved in transforming 1 g of ice at 0° and 1 atm into vapor at 150° C under 0.1 atm? Assume C_p of liquid water to be 1.0 cal/gram and C_p of water vapor 0.45 cal/gram. *Ans.* 2.23 eu

13. Eastman and McGavock [*J. Am. Chem. Soc.* **59**, 145 (1937)] list the atomic heat capacity of rhombic S from 15° K to 360° K. From the data of Table II in this article determine by a graphical method the entropy of rhombic S per mole at 298.2° K.

14. From the data of Table 2 find the entropy changes accompanying the following reactions at 25° C.

    (a) $CO(g) + 2\,H_2(g) = CH_3OH(l)$

    (b) $2\,HgCl(s) = 2\,Hg(l) + Cl_2(g)$

    (c) $MgO(s) + H_2(g) = H_2O(l) + Mg(s)$.
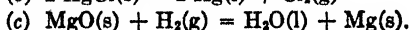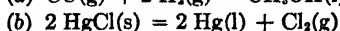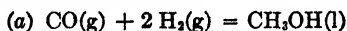
The Free Energy

Maximum Work and Tendency Toward Change. All changes in nature are due to the tendency on the part of various systems to reach a state of maximum stability commensurate with the state of each system. The state of maximum stability is the point of equilibrium. Once this state has been reached, the propensity toward further change disappears, and we say that the system is stable. As long as a system is away from equilibrium it will experience a tendency to reach that state, and the tendency will be greater the greater the distance from equilibrium.

Work results only when the tendency of systems to change toward equilibrium is harnessed in some way. From a system in equilibrium no work can be obtained, but a system on the way to equilibrium may be made to yield useful work. The amount of work that can be recovered from any system undergoing a change depends both on the nature of the change and the manner in which the system is harnessed. However, for each particular process there is a maximum amount of work the system can possibly do, and this maximum work may be taken as a measure of the tendency of the system in question to undergo change.

A system undergoing change can perform maximum work only when the change is carried out reversibly. If the process is not completely reversible, the amount of work obtainable is always below the maximum, the difference appearing as heat. The driving force behind the change is still, however, the maximum work difference between the final and initial states, for this difference still represents the highest possible quantity of energy that *tends* to appear in utilizable form as a result of the process. Whether it does appear or not depends entirely on the manner in which the process is conducted, and in no way affects the conclusion that the maximum work a process *may* perform, if conducted properly, is the true measure of the driving tendency behind the process.

Free Energy and Free Energy Change. The maximum work a process may yield is not necessarily the amount of energy available for doing useful work, even though the process is conducted reversibly. Of the total amount of work available, a certain amount has to be utilized for the performance of pressure-volume work against the atmosphere due

to contraction or expansion of the system during the process. For a process taking place at constant temperature and pressure, such as a chemical reaction at, say, 1 atmosphere pressure and $298°$ K, and involving a volume change from V_1 to V_2, the work done against the atmosphere is $P(V_2 - V_1) = P\Delta V$. Since this work is accomplished at the expense of the maximum work yielded by the process, the *net* amount of energy available for work other than pressure-volume against the confining atmosphere must be

$$\begin{aligned}
\text{Net available energy at } T \text{ and } P &= w_m - P\Delta V \\
&= -\Delta A - P\Delta V \\
&= -(\Delta A + P\Delta V)
\end{aligned} \tag{1}$$

To bring out more precisely the nature of the maximum net energy at constant temperature available from a process, let us define a new quantity F, called the *free energy* of a system, by the relation

$$F = A + PV \tag{2}$$

Since A, P, and V are characteristic of the state of the system only, F must also have this attribute; i.e., *F must also be a function characteristic of the state of the system only and be independent of the manner of arriving at that state.* The change in free energy between any two states will be, therefore,

$$\begin{aligned}
\Delta F &= F_2 - F_1 \\
&= (A_2 + P_2 V_2) - (A_1 + P_1 V_1) \\
&= \Delta A + (P_2 V_2 - P_1 V_1)
\end{aligned} \tag{3}$$

and at constant pressure

$$\begin{aligned}
\Delta F &= \Delta A + P(V_2 - V_1) \\
&= \Delta A + P\Delta V
\end{aligned} \tag{4}$$

If equation (4) is substituted in equation (1), we see that

$$\begin{aligned}
\text{Net available energy at } T \text{ and } P &= -(\Delta A + P\Delta V) \\
&= -\Delta F
\end{aligned} \tag{5}$$

This relation gives the significance of the free energy function. We see that the net available energy results from a *decrease* in the free energy content of the system in passing from the initial to the final state, and hence the free energy represents the maximum work content of a system corrected for the amount of energy tied up in the product PV.

The free energy change for any process, being a function of the initial and final states of the system only, is a definite quantity at any given temperature and pressure, and varies as these two variables are changed

As in the case of heat contents and internal energies, the absolute values of the free energies of substances are not known, and hence only differences can be dealt with. The free energy changes of processes are expressed in equations similar to thermochemical ones, and can be similarly added and subtracted. Thus, for instance, the free energy changes for the two reactions below are:

$$SO_2(g) + Cl_2(g) = SO_2Cl_2(g) \qquad \Delta F_{298° K} = -2,310 \text{ cal} \qquad (6a)$$
$$S(\text{rhom.}) + O_2(g) + Cl_2(g) = SO_2Cl_2(g) \quad \Delta F_{298° K} = -74,060 \text{ cal} \quad (6b)$$
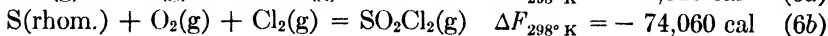
If equation (6b) is subtracted from (6a), we find that

$$SO_2(g) + Cl_2(g) - S(\text{rhom.}) + O_2(g) - Cl_2(g) = 0$$
$$\Delta F_{298° K} = +71,750 \text{ cal}$$

$$\text{or} \qquad S(\text{rhom.}) + O_2(g) = SO_2(g) \qquad \Delta F_{298° K} = -71,750 \text{ cal} \qquad (6c)$$
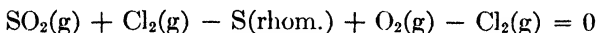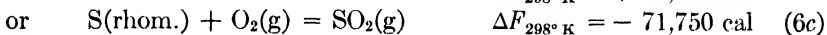
Consequently, the formation of sulfur dioxide from rhombic sulfur and gaseous oxygen proceeds with a free energy decrease of 71,750 calories; i.e., the free energy content of $(S + O_2)$ is greater than that of SO_2 by this amount.

Sign of the Free Energy Change. The sign of the free energy change of a process is very significant. When the driving tendency of a reaction

$$A + B = C + D$$
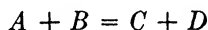
is from left to right, energy is emitted on reaction, and the sign of ΔF is negative. A minus sign denotes, therefore, that the reaction tends to proceed spontaneously. When the tendency is from right to left, however, net work equivalent to ΔF has to be absorbed in order for the reaction to proceed in the direction indicated, and ΔF is positive. A positive sign for ΔF signifies, therefore, that the reaction in the given direction is not spontaneous. Finally, when the system is in equilibrium, there is no tendency to proceed in either direction, no work can be done by the system, and hence $\Delta F = 0$. These three possible conditions for the free energy change of a process may be summarized as follows:

$$A + B \longrightarrow C + D \qquad \Delta F = - \text{ (spontaneous)} \qquad (7a)$$
$$A + B \longleftarrow C + D \qquad \Delta F = + \text{ (nonspontaneous)} \qquad (7b)$$
$$A + B \rightleftharpoons C + D \qquad \Delta F = 0 \text{ (equilibrium)} \qquad (7c)$$
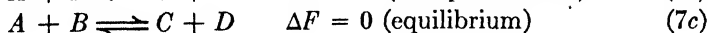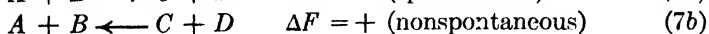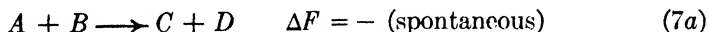
The arrows indicate the directions the reaction tends to follow spontaneously for the given sign of the free energy change.

A negative free energy change for a process does not necessarily mean that the process will take place. It is merely an indication that the process *can* occur provided the conditions are right. Thus oxygen and hydrogen in the ratio of 1 : 2 can coexist indefinitely at room temperature

without combining to form water although ΔF for the reaction at 25° C is $-$ 56,690 calories per mole of water. When a catalyst like platinized asbestos is introduced, however, the reaction proceeds with explosive violence. Even with a catalyst the reaction would have been impossible had not the potentiality to react been present. It is the sign of the free energy change which determines whether the potentiality to react exists, and it is the magnitude of the free energy change which tells us how large that potentiality is.

The Gibbs-Helmholtz Equation. The relationship between the free energy function and other thermodynamic quantities previously defined is readily deducible. If A is eliminated between equation (2) of Chapter X $(E = A + TS)$ and equation (2) of this chapter $(F = A + PV)$, we obtain a new defining expression for F, or

$$
\begin{aligned}
F = A + PV &= E - TS + PV \\
&= (E + PV) - TS \\
&= H - TS
\end{aligned} \tag{8}
$$

From equation (8) the free energy change at constant temperature for a process between two states, 1 and 2, follows as

$$
\begin{aligned}
\Delta F &= F_2 - F_1 \\
&= (H_2 - TS_2) - (H_1 - TS_1) \\
&= (H_2 - H_1) - T(S_2 - S_1) \\
&= \Delta H - T\Delta S
\end{aligned}
$$

or $\qquad \Delta F - \Delta H = - T\Delta S \qquad\qquad$ (9)

Equation (9) relates directly the changes in heat content, entropy, and free energy of any process. It is known as the *Gibbs-Helmholtz equation*. It is equivalent to equation (3) of Chapter X. If this equation is written as

$$
\Delta H = \Delta F + T\Delta S
$$

we see that the total change in heat content of the process at constant temperature and pressure is composed of (1) the net maximum work available from the process, ΔF, and (2) the total amount of energy that is unavailable for doing work, $T\Delta S$. It is the presence of the last term which eliminates the use of ΔH, although it does represent an energy gradient, as a measure of the driving tendency of the process. As may be seen from the equation, the unavailable energy $T\Delta S$ must be subtracted from the total energy change ΔH to obtain a measure of the driving tendency ΔF.

An alternate relation between ΔF and ΔH which is of considerable utility may be obtained from equation (9). If this equation be differ-

entiated with respect to T at constant P, then

$$\left[\frac{\partial(\Delta F)}{\partial T}\right]_P - \left[\frac{\partial(\Delta H)}{\partial T}\right]_P = -T\left[\frac{\partial(\Delta S)}{\partial T}\right]_P - \Delta S \qquad (10)$$

Since, however,

$$\left[\frac{\partial(\Delta H)}{\partial T}\right]_P = \Delta C_p \qquad \left[\frac{\partial(\Delta S)}{\partial T}\right]_P = \frac{\Delta C_p}{T}$$

equation (10) becomes

$$\left[\frac{\partial(\Delta F)}{\partial T}\right]_P - \Delta C_p = -\frac{T\Delta C_p}{T} - \Delta S$$

and therefore,

$$\left[\frac{\partial(\Delta F)}{\partial T}\right]_P = -\Delta S \qquad (11)$$

Substituting equation (11) in equation (9), the latter becomes

$$\Delta F - \Delta H = T\left[\frac{\partial(\Delta F)}{\partial T}\right]_P \qquad (12)$$

The physical significance of equations (9) or (12) is best illustrated by an example. Consider the reaction

$$Zn + CuSO_4 = ZnSO_4 + Cu \qquad (13)$$
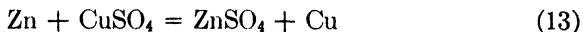
When this reaction is permitted to take place in an open beaker by adding zinc to a solution of copper sulfate, heat of reaction equal to ΔH is obtained. If, on the other hand, the same reaction is carried out reversibly by allowing the process to proceed in an electromotive cell, and forcing the voltage established to do work against another voltage only infinitesimally smaller than that of the cell, then instead of all heat, work equivalent to ΔF will be obtained. The difference between the work so obtained and the heat which would have been liberated had the reaction been carried out completely irreversibly, as in an open beaker, is given by the Gibbs-Helmholtz equation, and is equal to either $-T\Delta S$ or $T[\partial(\Delta F)/\partial T]_P$.

The term $-T\Delta S = T[\partial(\Delta F)/\partial T]_P$ represents the heat interchange between the system and its surroundings when the process is conducted *reversibly*, for

$$-T\Delta S = -T\frac{q}{T} = -q$$

and hence,

$$\Delta H - \Delta F = q \qquad (14)$$

When ΔH is greater than ΔF, the difference q is positive, and energy is absorbed as heat from the surroundings. On the other hand, when ΔF is greater than ΔH, q is negative, and heat is evolved to the surroundings. Finally, in the special case when $\Delta H = \Delta F$, heat is neither ab-

sorbed nor evolved by the system, and hence there is no change in entropy. Conversely, it must follow that only when there is no change in entropy on reaction will the maximum net work available from a process, ΔF, be equal the change in heat content, ΔH.

Variation of ΔF with Temperature. The free energy change of a process, like the heat of reaction, varies with temperature. The manner in which ΔF varies with temperature may be obtained by differentiating the quantity $\Delta F/T$ with respect to T at constant pressure. Then

$$\left[\frac{\partial\left(\frac{\Delta F}{T}\right)}{\partial T}\right]_P = \frac{T\left[\frac{\partial(\Delta F)}{\partial T}\right]_P - \Delta F}{T^2}$$

But, according to (12),

$$T\left[\frac{\partial(\Delta F)}{\partial T}\right]_P = \Delta F - \Delta H$$

Therefore,

$$\left[\frac{\partial\left(\frac{\Delta F}{T}\right)}{\partial T}\right]_P = \frac{\Delta F - \Delta H - \Delta F}{T^2}$$

$$= -\frac{\Delta H}{T^2} \tag{15}$$

Equation (15) gives the variation of ΔF with temperature in terms of ΔH, the heat of reaction, and may be employed to calculate ΔF at one temperature from that at another. To do this, ΔH as a function of temperature, deduced by the method outlined in Chapter IX, must be available, as well as one value of ΔF at a known temperature. With such information at hand, equation (15) can be integrated in the manner illustrated below.

Suppose ΔF is sought at 1000° K for the reaction

$$\frac{1}{2}\,N_2(g) + \frac{3}{2}\,H_2(g) \doteq NH_3(g) \qquad \Delta F_{298.2°\,K} = -3980\ \text{cal} \tag{16}$$
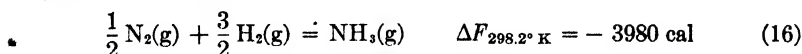
For this reaction it was already shown, p. 284, that ΔH is given by

$$\Delta H = -9560 - 6.06\,T + 2.19 \times 10^{-3}\,T^2 + \frac{0.398 \times 10^5}{T}$$

Substituting this value of ΔH in equation (15), we have

$$\left[\frac{\partial\left(\frac{\Delta F}{T}\right)}{\partial T}\right]_P = \frac{-\Delta H}{T^2} = \frac{9560}{T^2} + \frac{6.06}{T} - 2.19 \times 10^{-3} - \frac{0.398 \times 10^5}{T^3}$$

On integration this expression becomes

$$\frac{\Delta F}{T} = \frac{-9560}{T} + 6.06\ln T - 2.19 \times 10^{-3}\,T + \frac{0.199 \times 10^5}{T^2} + I$$

where I is a constant of integration. Hence,

$$\Delta F = -9560 + 6.06\,T \ln T - 2.19 \times 10^{-3}\,T^2 + \frac{0.199 \times 10^5}{T} + IT$$

Inserting now the value of $\Delta F = -3980$ calories at $T = 298.2°$ K from equation (16), and solving for I, we find that

$$-3980 = -9560 + 6.06(298.2) \ln (298.2) - 2.19 \times 10^{-3}(298.2)^2$$
$$+ \frac{0.199 \times 10^5}{298.2} + I(298.2)$$

$$I = -15.39$$

Consequently, ΔF as a function of temperature follows as

$$\Delta F = -9560 + 6.06\,T \ln T - 2.19 \times 10^{-3}\,T^2 + \frac{0.199 \times 10^5}{T} - 15.39\,T \quad (17)$$

and hence, at $T = 1000°$ K, ΔF is

$$\Delta F = -9560 + 6.06(1000) \ln (1000) - 2.19 \times 10^{-3}(1000)^2$$
$$+ \frac{0.199 \times 10^5}{1000} - 15.39(1000)$$

$$= 14,760 \text{ cal}$$

We see, therefore, that although the formation of ammonia will proceed spontaneously at 298° K, for ΔF is negative, the same process at 1000° K will be nonspontaneous to the extent of 14,760 calories.

The method described for calculating ΔF at one temperature from that at another is thermodynamically exact, and is conditioned only by the accuracy of the data employed. As in the case of heat of reaction, great care must be exercised that calculations made are within the range over which the thermal data have been determined. Only with precise data and with due regard for the limits of their validity can great trust be placed in the results of such calculations.

Other Applications of ΔF Equations. ΔF equations such as equation (17) are useful not only for calculating the free energy change at any temperature T, but also for evaluating ΔS and ΔH. ΔS is readily obtained by differentiation with respect to temperature at constant pressure of the ΔF expression in accordance with equation (11). In turn, ΔH may be obtained either by differentiating $\Delta F/T$ with respect to T and using equation (15), or by evaluating first both ΔF and ΔS, and subsequently employing the Gibbs-Helmholtz equation (9) to find ΔH.

Example: Suppose ΔS and ΔH for the ammonia synthesis, equation (16), are to be calculated at 1000° K. To obtain ΔS we differentiate first equation (17) with respect to T. Then,

$$\left[\frac{\partial(\Delta F)}{\partial T}\right]_P = -\Delta S = 6.06 + 6.06 \ln T - 4.38 \times 10^{-3}\,T - \frac{0.199 \times 10^5}{T^2} - 15.39$$

$$\Delta S = 9.33 - 6.06 \ln T + 4.38 \times 10^{-3}\,T + \frac{0.199 \times 10^5}{T^2}$$

On substituting $T = 1000°$ K, we find thus $\Delta S_{1000°\text{K}} = -28.15$ eu. Having now ΔF and ΔS at $1000°$ K, the Gibbs-Helmholtz equation yields for ΔH

$$\Delta H = T\Delta S + \Delta F$$
$$= 1000(-28.15) + 14,760$$
$$= -13,390 \text{ cal at } 1000° \text{ K}$$

This result is identical with the one found for the heat of this reaction at $1000°$ K on p. 284 in Chapter IX.

Free Energy and Pressure. The free energy of a substance varies not only with temperature but also with change in pressure. By a thermodynamic argument which need not be gone into here,[1] it can be shown that the variation of free energy with pressure at constant temperature is given by

$$\left(\frac{\partial F}{\partial P}\right)_T = V \tag{18}$$

where V is the volume of the substance at temperature T and pressure P. From equation (18) the change in free energy attending a change in total pressure on a substance from P_1 to P_2 at the constant temperature T follows as

$$\int_{F_1}^{F_2} dF = \int_{P_1}^{P_2} V dP$$
$$\Delta F = F_2 - F_1 = \int_{P_1}^{P_2} V dP \tag{19}$$

To evaluate the integral between the given limits, V must be known as a function of P.

The Free Energy of Ideal Substances. For ideal gases equation (19) is readily integrable. Since for an ideal gas $V = nRT/P$, equation (19) becomes

$$\Delta F = \int_{P_1}^{P_2} V dP$$
$$= \int_{P_1}^{P_2} \frac{nRT dP}{P}$$
$$= nRT \ln \frac{P_2}{P_1} \tag{20}$$

Equation (20) gives the change in free energy of n moles of an ideal gas due to change in pressure from P_1 to P_2 at the constant temperature T.

Equation (20) is also valid for the free energy change with change in osmotic pressure of ideal solutions obeying van't Hoff's law, $\Pi V = nRT$. When the osmotic pressure Π of an ideal solution containing n moles of

[1] See Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.

solute in V liters of solution is changed from Π_1 to Π_2, V is given by $V = nRT/\Pi$. The free energy change due to the process is again that given by equation (20), except that now P is replaced by the osmotic pressure Π, and n is the number of moles of solute in solution. Since, however, $\Pi = (n/V)RT = CRT$, where C is the concentration in moles per liter, then,

$$\frac{\Pi_2}{\Pi_1} = \frac{C_2 RT}{C_1 RT} = \frac{C_2}{C_1}$$

and equation (20) becomes for ideal solutions

$$\Delta F = nRT \ln \frac{C_2}{C_1} \tag{21}$$

Equation (21) gives the free energy change involved in transferring n moles of solute from a solution of concentration C_1 to a solution of concentration C_2.

As solids and liquids are only slightly compressible, their volume may be considered essentially constant over appreciable pressure ranges. Equation (20) integrates, then, simply to

$$\Delta F = \int_{P_1}^{P_2} V dP$$
$$= V(P_2 - P_1) \tag{22}$$

where ΔF is the free energy change of the solid or liquid due to change in pressure from P_1 to P_2. Such free energy changes with pressure usually are negligible compared to the free energy changes in gases and solutions, and may be disregarded; i.e., the free energies of pure solids and liquids may be considered to be constant over a fairly wide pressure range at any given temperature.

The Fugacity and Activity Concepts. When equation (20) is applied either to real gases at high pressures or to nonideal solutions, especially those containing electrolytes even at low concentrations, it is found that the change in free energy is not reproduced by this simple relation. The difficulty is, of course, that in cases of nonideal behavior V is no longer given by nRT/P, but by some more complicated function of the pressure. To obtain the free energy change with pressure for the nonideal substance, the exact dependence of the volume on the pressure must be known before equation (19) can be integrated. Since such volume dependence may be highly individual, the result would be that ΔF in each case would be given by an equation different in form, and the simplicity and generality of equation (20) would be destroyed.

G. N. Lewis first showed how nonideal systems may be handled without discarding the simple free energy equations deduced for ideal systems.

To do this he introduced two new thermodynamic quantities, *fugacity* and *activity*. To understand these quantities, consider first a system composed of liquid water and its vapor. At constant temperature there is a definite pressure of water vapor above the liquid. This vapor comes from the liquid phase, and represents a tendency of the liquid to pass into the vapor phase. In turn, the vapor tends to escape the gaseous state by condensing to liquid. When these two *escaping tendencies* become equal, the system reaches equilibrium, i.e., the vapor pressure becomes constant at constant temperature. We may say, therefore, that a state of equilibrium is the point at which the escaping tendency of a constituent is the same in all parts of the system, and not the point at which the escaping tendency is zero.

The idea that each substance in a particular state has a definite tendency to escape from that state is perfectly general. Lewis pointed out that this escaping tendency can be measured by a quantity f, called the *fugacity*, which is related to the free energy content of the substance per mole, F, by the expression

$$F = RT \ln f + B \qquad (23)$$

where B is a constant dependent only on the temperature and the nature of the substance. Since absolute values of the free energy are not known, B cannot be evaluated. However, we can get around this difficulty by referring all free energy measurements to a reference point. If we designate by F^0 the free energy and by f^0 the fugacity in this reference state, then F^0 is given by

$$F^0 = RT \ln f^0 + B \qquad (24)$$

and the free energy difference between any state in which the free energy is F and the reference state is given by

$$F - F^0 = RT \ln \frac{f}{f^0} \qquad (25)$$

Consequently, the free energy content of a substance in any state in terms of the free energy in the reference state follows as

$$F = F^0 + RT \ln \frac{f}{f^0} \qquad (26)$$

If we write now

$$\frac{f}{f^0} = a \qquad (27)$$

equation (26) becomes

$$F = F^0 + RT \ln a \qquad (28)$$

The quantity a is called the *activity*. From equation (28) we see that the free energy per mole of any substance at a temperature T may be written as the free energy of the substance in a reference state at the

same temperature, and the quantity $RT \ln a$. In the reference state $F = F^0$, $RT \ln a = 0$, and hence $a = 1$; i.e., in the reference state the activity must be unity. In any other state the value of the activity will depend on the difference $(F - F^0)$, or in other words, on the distance of the particular state from the reference state. This relation between F and F^0 can best be understood from Fig. 1. Let the vertical line represent a free energy axis. On this axis A represents the arbitrary point chosen for reference, while B represents another point where the free energy is F_2. Then equation (28) states that the difference in free energies of a substance between points B and A is $RT \ln a_2$. Similarly the free energy difference between a point like C and the reference state is given by $RT \ln a_1$.
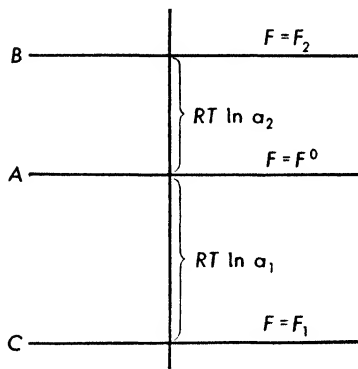
Fig. 1

In terms of equation (28) the difference in free energy per mole occasioned on passing from one state where the free energy is F_1 and the activity is a_1, to another state where the free energy is F_2 and the activity is a_2, must be

$$\Delta F = F_2 - F_1 = (F^0 + RT \ln a_2) - (F^0 + RT \ln a_1)$$

$$\Delta F = RT \ln \frac{a_2}{a_1} \tag{29a}$$

or, for n moles,

$$\Delta F = nRT \ln \frac{a_2}{a_1} \tag{29b}$$

The striking similarity of equations (29a) and (29b) to equations (20) and (21) suggests that we may consider the activity as the thermodynamic counterpart of the concentration or pressure. When activities of gases and solutions, as determined according to well-worked-out thermodynamic methods, are substituted for pressures and concentrations in equations (20) and (21), exact free energy calculations are possible. The reason for this exactness is that evaluation of activities, unlike that of concentrations and pressures, takes into account not only stoichiometric relationships, but also any mutual·attractions between molecules, interactions between solute and solvent in a solution, and ionization. These effects complicate ideal behavior, and are the factors responsible for the breakdown of the thermodynamic equations for ideal systems when applied to real substances.

Standard States. The particular state of a substance in which $a = 1$ *is called the standard state of the substance.* The standard state is, according to equation (28), the state in which $F = F^0$; i.e., the state with reference to which all free energy calculations are to be made. Before activities of gases, liquids, solids, and solutions can be evaluated, the particular state to be chosen as a reference must be defined. This definition may be purely arbitrary, but certain established conventions are in general use at present. These will be discussed now.

Standard States for Gases. As the standard state for a gas at any given temperature is taken the state in which the fugacity is equal to unity, namely, $f^0 = 1$. On the basis of this definition, the activity of any gas becomes equal to the fugacity, for

$$a = \frac{f}{f^0} = \frac{f}{1} = f \tag{30}$$

and hence for a gas equation (28) may be written as

$$F = F^0 + RT \ln f \tag{31}$$

For an ideal gas the fugacity is equal to the pressure. Since, however, any gas can be brought into an ideal state by reducing its pressure to zero, we can complete the definition of the fugacity of any gas by stating that, in general,

or that

$$f = P \text{ as } P \longrightarrow 0 \tag{32a}$$

$$\frac{f}{P} = 1 \text{ as } P \longrightarrow 0 \tag{32b}$$

As long as a gas is ideal the ratio f/P remains equal to unity. However, as soon as a gas deviates from ideal behavior, f is no longer equal to P, and the ratio f/P becomes something other than one. The further this ratio is from unity the greater is the nonideality. Consequently, this ratio, called the *activity coefficient* of a gas, and represented by the symbol γ, gives a direct measure of the extent to which any real gas deviates from ideality at any given pressure and temperature.

The methods employed evaluating fugacities of gases at various temperatures and pressures are beyond the scope of this book; they are more the concern of textbooks on chemical thermodynamics. Still, the general conclusions deduced from such calculations may be summarized as follows. For all gases at relatively low pressures, pressure may be substituted for fugacity in thermodynamic equations without the introduction of any serious error. The exact pressure up to which this is permissible cannot be specified, for it will depend on the nature of the gas in question, the temperature, and the accuracy required. At higher pressures, however, such a substitution may lead to considerable, and even very

large, errors, as may be judged from Table 1, where the fugacities and activity coefficients at various pressures are shown for the two gases, oxygen and carbon dioxide. An idea of the error that may be involved in the substitution of pressure for fugacity may be gathered from the following example.

TABLE 1

RELATION BETWEEN FUGACITY AND PRESSURE FOR SOME GASES [1]

P (atm)	f (atm)	$\gamma = f/P$
O_2 at $0°$ C		
50	48.0	0.960
100	92.5	0.925
200	174.	0.87
400	338.	0.85
600	540.	0.90
CO_2 at $60°$ C		
25	23.2	0.928
50	42.8	0.86
100	70.4	0.70
200	91.	0.45
300	112.	0.37

Example: Calculate the free energy change accompanying the compression of 1 mole of CO_2 at $60°$ C from 25 to 300 atm. Using pressures first, we obtain, according to equation (20), for ΔF

$$\Delta F = nRT \ln \frac{P_2}{P_1}$$

$$= 1 \times 1.987 \times 333 \times 2.303 \log_{10} \frac{300}{25}$$

$$= 1640 \text{ cal}$$

Using, instead, the fugacities as given in Table 1, ΔF follows from equation (29b) as

$$\Delta F = nRT \ln \frac{a_2}{a_1} = nRT \ln \frac{f_2}{f_1}$$

$$= 1 \times 1.987 \times 333 \times 2.303 \log_{10} \frac{112}{23.2}$$

$$= 1040 \text{ cal}$$

The approximate calculation gives here, therefore, a result 57.7% too high.

Standard States for Solids and Liquids. *As the standard state of a pure solid or a pure liquid is taken the solid or liquid at 1 atmosphere pressure at each temperature.* In this state of the solid or liquid $a = 1$ and $F = F^0$. Since, as we have seen, the free energies of solids and liquids are essentially independent of pressure, the activities must also remain constant, and hence for solids and liquids $a = 1$, to a very near approximation, at all temperatures and for wide ranges of pressures.

[1] Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923, pp. 196–197.

Standard States for Components of a Solution. The activity of each component of a solution is unity in the standard state of the given component, but the particular concentration of a solution that corresponds to each of these standard states is different. This difference is due to the manner in which the standard states of solute and solvent are defined. As the standard state of the solvent is chosen the *pure solvent.* On the other hand, as the standard state of the solute is taken the concentration of solution in which $a = 1$, the activity being evaluated on the basis that $a = C$ as $C \longrightarrow 0$, C being the concentration of solute in solution. In other words, in the infinitely dilute solution the activity of solute is set equal to the concentration; then, whatever concentration of solution contains solute at unit activity is the defined standard state of the solute.

Activities of dissolved substances may be determined from vapor pressure, freezing point lowering, solubility, electromotive force, and other data. Some of the methods involved are quite complicated, others are relatively simple. At this point it may be of interest to know that for a solution whose solvent vapor pressure is P_1, the activity of the *solvent in solution* is given by

$$a_1 = \frac{f_1}{f_1^0} = \frac{P_1}{P_1^0} \tag{33}$$

where P_1^0 is the vapor pressure of the pure solvent. This equation is valid provided the vapor at the pressures in question behaves ideally.

The Reaction Isotherm. With the free energy of a substance defined in terms of the free energy in the standard state and the activity, we may proceed to deduce the equation for the free energy change involved in a chemical reaction. For a reaction such as

$$aA + bB + \cdots = cC + dD + \cdots \tag{34}$$
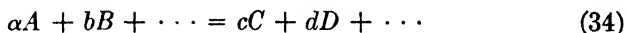
if the activities of A and B at the *start* are a_A and a_B, while the activities of C and D at the *end* of the reaction are a_C and a_D respectively, then the free energies of each of these substances *per mole* at a temperature T are given by the expressions:

$$F_A = F_A^0 + RT \ln a_A \tag{35a}$$
$$F_B = F_B^0 + RT \ln a_B \tag{35b}$$
$$F_C = F_C^0 + RT \ln a_C \tag{35c}$$
$$F_D = F_D^0 + RT \ln a_D \tag{35d}$$

where the F^0's with the given subscripts represent the free energies at unit activity of the respective species. By definition the free energy change of the reaction, ΔF, is

$$\Delta F = (cF_C + dF_D) - (aF_A + bF_B)$$

and hence,

$$\Delta F = (cF_C^0 + cRT \ln a_C + dF_D^0 + dRT \ln a_D)$$
$$- (\alpha F_A^0 + \alpha RT \ln a_A + bF_B^0 + bRT \ln a_B)$$
$$= [(cF_C^0 + dF_D^0) - (\alpha F_A^0 + bF_B^0)]$$
$$+ [(cRT \ln a_C + dRT \ln a_D) - (\alpha RT \ln a_A + bRT \ln a_B)]$$
$$= [(cF_C^0 + dF_D^0) - (\alpha F_A^0 + bF_B^0)] + RT \ln \frac{a_C^c a_D^d}{a_A^\alpha a_B^b} \tag{36}$$

The first term on the right of equation (36) gives the free energy change of the reaction in the standard state; i.e., the free energy change involved on reaction when the starting materials at *unit* activity react to form products also at *unit* activity. If this free energy change in the standard state be designated by ΔF^0, then equation (36) becomes

$$\Delta F = \Delta F^0 + RT \ln \frac{a_C^c a_D^d}{a_A^\alpha a_B^b} \tag{37}$$

Equation (37) gives the free energy change of a reaction as a function of the free energy change in the standard state and the starting and final activities at any *constant temperature T*. For this reason this most fundamental thermodynamic equation is referred to as the *reaction isotherm*. When the activities of the starting and final materials are all one, the second term on the right of equation (37) is zero, and $\Delta F = \Delta F^0$. On the other hand, when the initial and final activities are other than unity, the second term is not zero, and ΔF is given then by the full equation (37).

For any reaction ΔF^0 is constant at any given temperature and is completely independent of pressure. The variation of ΔF^0 with temperature, like that of ΔF, is given by equations (11) or (15), except that in this case the ΔH's and ΔS's are ΔH^0's and ΔS^0's; i.e., the heats of reaction and entropy changes are for the reactions in the standard states. Outside of this one restriction, calculations involving ΔF^0's are handled in exactly the same manner as the calculations involving ΔF's described in this chapter.

Standard Free Energies of Formation. Because of the direct relation of ΔF^0 to the conditions prevailing at equilibrium, to be discussed in the next chapter, the evaluation of the standard free energy change of a reaction is a problem of prime concern in physical chemistry. Free energies may be obtained from the heat of reaction and the entropy change at a particular temperature through the Gibbs-Helmholtz equation (9) and by other methods to be developed further in the text. At this time will be presented a scheme employed to tabulate free energies of formation of compounds, from which free energies of all types of reactions may be obtained.

The standard free energy of formation of a compound is defined as the free energy change accompanying the formation of the compound at unit activity from the elements also at unit activity. The equation

$$H_2(g) + \frac{1}{2}O_2(g) = H_2O(l) \qquad \Delta F^0_{298° K} = -56,690 \text{ cal} \qquad (38)$$
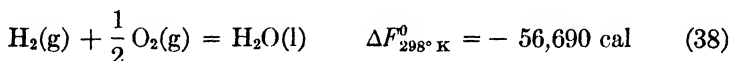
gives the change in free energy on formation of 1 mole of $H_2O(l)$ at unit activity from $H_2(g)$ and $1/2 \, O_2(g)$, both at unit activity. *Assuming now arbitrarily*, as in heats of formation, *that the free energies of formation of the elements in their standard states at 25° C are zero*, $\Delta F^0_{298° K}$ above becomes the standard free energy of formation of liquid water from the elements, and we may write simply that

$$H_2O(l) : \qquad \Delta F^0_{298° K} = -56,690 \text{ cal} \qquad (39)$$
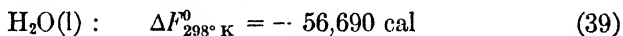
Table 2 lists the standard free energies of formation per mole for a number of compounds at 25° C. These free energies of formation may be used to calculate ΔF^0 of reactions in the same manner that heats of formation are used to calculate heats of reaction. Thus, for instance, taking the data from the table mentioned, we find at 25° C for the standard free energy change of the reaction

$$NO(g) + \frac{1}{2}O_2(g) = NO_2(g)$$

$$\Delta F^0_{25° C} = \Delta F^0_{NO_2} - \left[\Delta F^0_{NO} + \frac{1}{2}\Delta F^0_{O_2} \right]$$

$$= 12,270 - 20,660 - 0$$

$$= -8390 \text{ cal}$$
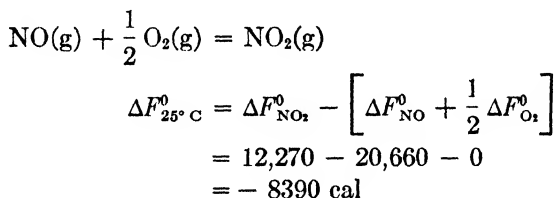
TABLE 2

STANDARD FREE ENERGIES OF FORMATION AT 25° C

Substance	ΔF^0 (cal/mole)	Substance	ΔF^0 (cal/mole)
$HCl(g)$	$-22,690$	$Ag_2O(s)$	$-2,590$
$H_2O(l)$	$-56,690$	$HgO(s)$	$-13,940$
$H_2O(g)$	$-54,670$	$NH_4OH(aq)$	$-62,990$
$CO(g)$	$-33,010$	$CH_4(g)$	$-12,280$
$CO_2(g)$	$-94,450$	$C_2H_6(g)$	$-10,700$
$NO(g)$	$20,660$	$C_2H_4(g)$	$12,300$
$NO_2(g)$	$12,270$	$C_2H_2(g)$	$50,750$
$N_2O_4(g)$	$23,440$	$C_6H_6(g)$	$30,640$
$H_2S(g)$	$-7,870$	$C_6H_6(l)$	$29,400$
$SO_2(g)$	$-71,740$	$CH_3OH(l)$	$-39,960$
$NH_3(g)$	$-3,980$	$C_2H_5OH(l)$	$-40,200$
$NaCl(s)$	$-91,700$	$HCOOH(l)$	$-85,200$
$AgCl(s)$	$-26,220$	$CH_3COOH(l)$	$-94,500$

REFERENCES FOR FURTHER READING

1. J. A. V. Butler, "Fundamentals of Chemical Thermodynamics," The Macmillan Company, New York, 1934–1935, 2 vols.
2. Hougen and Watson, "Industrial Chemical Calculations," second edition, John Wiley & Sons, Inc., New York, 1936.
3. Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.
4. F. H. MacDougall, "Thermodynamics and Chemistry," third edition, John Wiley & Sons, Inc., New York, 1939.
5. D. J. Martin, "Thermodynamics for Chemists," Longmans, Green and Company, New York, 1933.
6. Parks and Huffman, "The Free Energies of Organic Compounds," Reinhold Publishing Corporation, New York, 1932.

PROBLEMS

1. For the reaction, H_2 (1 atm) $+ 1/2\ O_2$ (1 atm) $= H_2O(l)$, calculate the difference in calories between ΔF and ΔA at 25° C.   *Ans.* $\Delta F - \Delta A = -888$ cal

2. From the following series of reactions and the free energy changes accompanying them find the free energy of formation of $N_2O_4(g)$ at 25° C:

$$\frac{1}{2}\ N_2(g) + \frac{1}{2}\ O_2(g) = NO(g) \qquad \Delta F^0_{298°\,K} = 20{,}660\ \text{cal}$$

$$NO(g) + \frac{1}{2}\ O_2(g) = NO_2(g) \qquad \Delta F^0_{298°\,K} = -8390\ \text{cal}$$

$$2\ NO_2(g) = N_2O_4(g) \qquad \Delta F^0_{298°\,K} = -1100\ \text{cal}$$
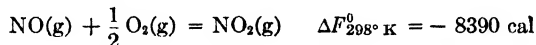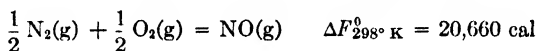
Ans. 23,440 cal/mole

3. From the appropriate absolute entropies and heats of formation found in the tables of the two preceding chapters calculate the free energy change at 25° C for the reaction:

$$CO(g) + H_2O(l) = CO_2(g) + H_2(g)$$
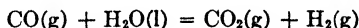
4. For the sublimation

$$Au(s) = Au(g)$$
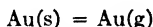
$$\Delta H^0_{298.2°\,K} = 90{,}500\ \text{cal/mole} \qquad \Delta F^0_{298.2°\,K} = 81{,}000\ \text{cal/mole}$$
$$Au(g): \quad C_p = 5.00\ \text{cal/mole}$$
$$Au(s): \quad C_p = 5.61 + 1.44 \times 10^{-3}\ T\ \text{cal/mole}$$

From these data find an expression for ΔF^0 as a function of T, and calculate ΔF^0 at 1000° K.

Ans. $\Delta F^0 = 90{,}740 + 1.40\ T \log_{10} T - 36.23\ T + 0.72 \times 10^{-3}\ T^2$

$$\Delta F^0_{1000°\,K} = 59{,}430\ \text{cal}$$

5. From the following data for the reaction,

$$\frac{1}{2}\ H_2(g) + \frac{1}{2}\ Cl_2(g) = HCl(g)$$
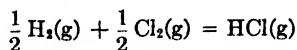
$$\Delta H^0_{291°\,K} = -22{,}000\ \text{cal} \qquad \Delta F^0_{298.1°\,K} = -22{,}700\ \text{cal}$$
and,
$$H_2(g): \quad C_p = 6.50 + 0.0009\ T\ \text{cal/mole}$$
$$Cl_2(g): \quad C_p = 7.4 + 0.001\ T\ \text{cal/mole}$$
$$HCl(g): \quad C_p = 6.50 + 0.001\ T\ \text{cal/mole}$$

deduce an expression for ΔF^0 as a function of T.

6. For the reaction, $FeCO_3(s) = FeO(s) + CO_2(g)$

$$\Delta F^0 = 18,660 - 14.42\, T \log_{10} T - 6.07\, T + 8.24 \times 10^{-3} T^2$$

Find ΔH^0 and ΔS^0 for the reaction at 25° C.

7. Assuming N_2 to be an ideal gas, find the free energy change involved in compressing 7 g of N_2 at 27° C from a pressure of 0.5 to 3 atm.   *Ans.* 267 cal

8. Calculate the free energy change for the transfer of 1 mole of naphthalene from a 0.01 molar solution in C_6H_6 to a 0.05 molar solution in the same solvent at 27° C. What is the significance of the sign of ΔF?

9. From the data of Table 1 find the free energy change accompanying the isothermal compression of 1 mole of O_2 from 50 to 600 atm at 0° C. Compare this value with the value which would have been obtained upon the assumption that O_2 behaves as an ideal gas.

10. For the reaction, $H_2(g, 1\ atm) + Cl_2(g, 1\ atm) = 2\,HCl(g, 1\ atm)$ ΔF^0 at 25° C is $-45,400$ cal. Find ΔF for the process

$$H_2(g, 2\ atm) + Cl_2(g, 1\ atm) = 2\,HCl(g, 0.1\ atm)$$
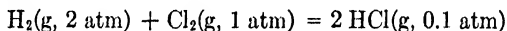
Ans. $-48,540$ cal

11. The standard free energy of formation of $H_2O(l)$ at 25° C is $-56,690$ cal. Find the free energy of formation of H_2O at 25° C from H_2 at a partial pressure of 0.01 atm and O_2 at a partial pressure of 0.25 atm.

12. From the data of Table 2, calculate ΔF^0 for the following reactions at 25° C:

   (a) $C_2H_5OH(l) + O_2(g) = CH_3COOH(l) + H_2O(l)$

   (b) $2\,CO_2(g) = 2\,CO(g) + O_2(g)$

   (c) $NH_3(g) + H_2O(l) = NH_4OH(aq)$

Which of the above reactions as written are spontaneous in the standard state?

13. Using the absolute entropies from Table 2 of the preceding chapter, calculate ΔH at 25° C for reactions (a) and (b) of the preceding problem.

14. At high temperatures ΔF^0 for the dissociation

$$CaCO_3(s) = CaO(s) + CO_2(g)$$

is given by the expression:

$$\Delta F^0 = 42,500 - 1.52\, T \log_{10} T - 34.6\, T + 2.15 \times 10^{-3} T^2$$

Calculate ΔF^0 at 100° intervals between 700 and 1200° C, plot the values as a function of T, and determine from the curve the temperature at which ΔF^0 changes sign.

Homogeneous Chemical Equilibrium

Introduction. In the study of chemical reactions per se, or in the development of chemical processes for industrial purposes, a knowledge of the laws governing chemical reactions is of inestimable value both to the chemist and the chemical engineer. Without theoretical guidance, the intelligent planning and direction of preliminary experimental and development work on any chemical process would be extremely difficult, and would require the experimental determination of a tremendous amount of complete data before final plant designs and operating procedures could be correctly defined. With the aid of theoretical principles, however, the amount of experimental work necessary for the definition of a process or reaction can be reduced to the bare minimum necessary for the application of the various established thermodynamic and kinetic equations.

An important question to be answered about any chemical reaction is: To what extent can the reaction in question take place? To answer this question completely it is necessary to ascertain the effects of temperature, pressure, foreign substances, and the state of the substances involved in the reaction, i.e., whether the reaction can take place better in the gaseous, liquid, or solid state, or in solution. The problem involved here is the one of chemical equilibrium, and is purely thermodynamic in nature. It can frequently be answered by a study of the free energy relations involved, provided the necessary data are available. This is the problem that will concern us in this and the following chapter.

Chemical Equilibrium. It is a familiar and well-established fact that many reactions do not go to completion. They proceed to a certain point and then apparently stop, often leaving considerable amounts of unaffected reactants. The remarkable fact is that, under any given set of conditions of temperature, pressure, and concentration, the point at which any reaction seems to stop is always the same; i.e., there exists at this point among the concentrations of the various reactants and products of any reaction a relationship which is definitely fixed. When a reaction reaches this stage in its course, it is said to be in *equilibrium*.

A state of chemical equilibrium should not be considered the state of a reaction at which all motion ceases. It is much more fruitful to consider the point of equilibrium as the state in which the *rate* at which reactants disappear to form products is exactly equal to the *rate* at which the products interact to reform the reacting substances. Under these conditions no perceptible transformation can be detected in the system, and the net effect is an *apparent* state of complete rest. Such an equilibrium is designated as a *dynamic* one, in contrast to a *static* equilibrium where there is no motion whatsoever. All chemical equilibria, as well as physical equilibria between states, are considered to be dynamic in nature.

In view of the concept of dynamic equilibrium, every reaction may be considered to be reversible and have a point of equilibrium. In reactions going spontaneously to completion, the state of equilibrium is far on the side of the products, the quantities of unreacted materials being so small as to escape detection by ordinary laboratory methods. On the other hand, in reactions designated as not taking place, the equilibrium is almost completely in favor of the reactants, the concentrations of products being beyond detection. The fact that we may not be able to detect the presence of reactants in one case or products in the other does not militate against the concept of dynamic equilibrium; for it is generally believed that the extremely small concentrations of substances demanded by some equilibria could be found if more sensitive and precise methods were at our disposal. In all instances where the principle has been tested it has been found to be valid in every detail.

Chemical equilibria may be classified into two groups: (1) *homogeneous* equilibria, and (2) *heterogeneous* equilibria. A homogeneous equilibrium is one established in a system in which only one phase occurs, as in a system containing only gases, or a single liquid or solid phase. A heterogeneous equilibrium, on the other hand, is one established in a system in which more than a single phase appears, as equilibrium between solid and gas, liquid and gas, solid and liquid, or solid and solid.

The Law of Mass Action. Although attempts had been made by various investigators to express chemical reactivity or affinity, up to about 1777 the effect of concentration in chemical reactions was not appreciated. In that year Wenzel first called attention to the importance of considering the relative amounts of reagents in discussing chemical reactions. It was de Berthollet, however, who first clearly pointed out the real role concentration may play in determining the course of a reaction, and this point was further emphasized by the important quantitative researches of Berthelot and Pean de St. Gilles.[1] In studying the reaction between alcohol and organic acids to form esters and water, the

[1] Berthelot and Pean de St. Gilles, Ann. chim. phys., **65**, 385; **66**, 5; **68**, 225 (1862–1863).

latter investigators found that the reactions were reversible and that the percentage of acid converted to ester depended on the relative amounts of alcohol and acid used: the greater the number of equivalents of alcohol per equivalent of acid, the greater was the percentage of acid converted to ester.

The various experimental results finally crystallized into a principle of utmost importance when the two Scandinavian investigators, Guldberg and Waage, enunciated in 1867 their celebrated *law of mass action*. The law of mass action states that at constant temperature *the rate of a chemical reaction is proportional to the "active masses" of the reacting substances present at any given time*. The ambiguous term in the above statement is "active mass." Guldberg and Waage appreciated this fact, and extended their statement to indicate that for gases and solutions "concentration" may be substituted for "active mass." For solids and pure liquids the "active mass" may be considered to be constant, since the rate of reaction is independent of the amounts of these participating in the reaction.

Mathematical Formulation of Law of Mass Action. With the aid of this principle may now be defined the relations which obtain among the concentrations of reactants and products at equilibrium. According to the law of mass action, the rate of the direct reaction v_1 in the reversible process,

$$A + B \cdots = C + D \cdots \tag{1}$$

must at any given time be proportional to C_A, the concentration of A, C_B, the concentration of B, etc. Therefore,

$$v_1 = k_1 C_A C_B \cdots \tag{2}$$

where k_1 is a constant of proportionality known as the *specific rate* constant for the direct reaction. k_1 represents the rate of the reaction when the concentrations of the reactants are all unity. Similarly v_2, the rate of the reverse reaction, is given at any instant by the expression

$$v_2 = k_2 C_C C_D \cdots \tag{3}$$

where k_2 is the specific velocity constant for the reverse reaction. Since at equilibrium the two rates must be equal, then for the condition of chemical equilibrium we have that

$$v_1 = v_2$$
$$k_1 C_A C_B \cdots = k_2 C_C C_D \cdots$$

and
$$\frac{C_C C_D \cdots}{C_A C_B \cdots} = \frac{k_1}{k_2} = K_c \tag{4}$$

According to equation (4) the ratio of the product of the concentrations of resultants to the product of the concentrations of reactants is equal at equilibrium to the ratio of the two velocity constants, and hence must be equal to another constant K_c. K_c is known as the *equilibrium constant* of the particular reaction. The subscript c is used to indicate that the ratio in equation (4) is expressed in concentration units, i.e., moles per liter.

When more than a single molecule of a substance is involved in a reaction, the rate is no longer proportional to the first power of the concentration of the reacting specie, but to the *concentration of the given substance raised to the power of the number of molecules of the substance participating*. Thus, for the reaction

$$aA + bB + \cdots = cC + dD + \cdots \qquad (5)$$
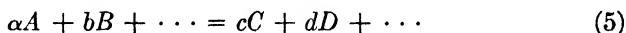
the rates are given by

$$v_1 = k_1 C_A^a C_B^b \cdots \qquad (6)$$
$$v_2 = k_2 C_C^c C_D^d \cdots \qquad (7)$$

and, since at equilibrium $v_1 = v_2$,

$$k_1 C_A^a C_B^b \cdots = k_2 C_C^c C_D^d \cdots$$
$$\frac{C_C^c C_D^d \cdots}{C_A^a C_B^b \cdots} = \frac{k_1}{k_2} = K_c \qquad (8)$$

Equation (8) gives the most general definition of the equilibrium constant. By convention the *concentrations of products are always placed in the numerator of the equilibrium constant expression*, the *concentrations of reactants* in the *denominator*. To obtain the correct equilibrium constant for a given reaction this convention must be adhered to rigidly. Reversal of the ratio will not give the constant for the reaction as written, but its reverse; for, as may be seen from equation (8), the two constants are reciprocally related to each other, i.e.

$$K_{\text{direct reaction}} = \frac{1}{K_{\text{reverse reaction}}} \qquad (9)$$

K_p and K_c. For reactions involving gases, it is usually more convenient to express the concentration of gas at any given temperature in terms of its partial pressure. If we represent by P_A, P_B, P_C, and P_D the partial pressures of the respective gaseous species present at equilibrium, then, for an equilibrium such as given in equation (5), we may write

$$\frac{P_C^c P_D^d \cdots}{P_A^a P_B^b \cdots} = \frac{k_1'}{k_2'} = K_p \qquad (10)$$

where K_p is the equilibrium constant in terms of partial pressures. The pressures may be expressed in any desired units, usually atmospheres.

In general, K_p and K_c for a gaseous reaction are not the same numerically. A relation between these two constants at any temperature T may readily be deduced provided the gases involved may be considered to behave ideally. Since for an ideal gas $P = (n/V)RT = CRT$, we may write, on substitution of this relation in equation (10),

$$\frac{C_C^c (RT)^c C_D^d (RT)^d}{C_A^a (RT)^a C_B^b (RT)^b} = K_p$$

$$\left(\frac{C_C^c C_D^d}{C_A^a C_B^b} \right) \frac{(RT)^{c+d}}{(RT)^{a+b}} = K_p$$

or
$$K_p = K_c (RT)^{(c+d)-(a+b)} \tag{11a}$$

But, $(c + d) - (a + b)$ represents the change in the total number of moles of gaseous products and reactants during the reaction. Letting this difference be Δn, we obtain for the relation between the two constants

$$K_p = K_c (RT)^{\Delta n} \tag{11b}$$

It is apparent from equation (11) that $K_p = K_c$ only when $\Delta n = 0$; i.e., when there is no change in volume on reaction. However, when there is a volume change, $K_p \neq K_c$. For an increase in volume on reaction Δn is positive, and hence K_p is numerically larger than K_c; on the other hand, when Δn is negative, corresponding to a volume decrease, K_p is less than K_c.

This relation between K_p and K_c may be illustrated by the following example. For the reaction

$$2 \, H_2S(g) = 2 \, H_2(g) + S_2(g)$$

K_p, expressed in atmospheres, is 0.0260 at 1132° C. To find K_c for concentrations in *moles per liter* we have $\Delta n = 3 - 2 = 1$, $R = 0.0821$ for liter-atmospheres, and $T = 1132 + 273 = 1405°$ K. Hence,

$$K_c = \frac{K_p}{(RT)^{\Delta n}} = \frac{0.0260}{(0.0821 \times 1405)} = 0.000225$$

R must be expressed in the same units as are the pressures and volumes involved in K_p and K_c.

Properties of Equilibrium Constants. Because of the fundamental importance of equilibrium calculations in physical chemistry and chemical engineering, it may not be amiss to recapitulate and emphasize the properties of the equilibrium constant of a reaction. For it is only through a thorough appreciation of the attributes of this relation that calculations and predictions can be made with certainty.

First of all, the equilibrium constant principle is valid only at *equilibrium*. The relation called for between the concentrations of products and reactants does not apply to every and all possible concentrations that may be encountered in a reacting system, but only to those concentrations which correspond to the point of true equilibrium. Unless true equilibrium concentrations are substituted in equations (8) or (10), these equations cannot be expected to yield constants.

Again, if the mass law and its formulation are correct, the equilibrium constant for any reaction should, for any given temperature, be a constant independent of concentration or pressure at all concentrations and pressures. A deviation from constancy outside the limits of experimental error may indicate either that the mass law is invalid, or that the definition of "active mass" as given by Guldberg and Waage is incorrect. On the other hand, the equilibrium constant of a reaction will not be the same at all temperatures, but will vary from temperature to temperature in a manner predictable by thermodynamics.

The magnitude of the equilibrium constant determines the extent to which any particular reaction can proceed under given conditions. A large value of K_p or K_c indicates that the numerator in the equilibrium constant expression is large compared to the denominator, i.e., that the concentrations of products are large compared to those of the reactants, and hence that the reaction favors the formation of products. On the other hand, when K_p or K_c is small, the concentrations of reactants are large compared to those of products, and the indications are that the particular reaction does not proceed to any appreciable extent under the given conditions.

Further, the equilibrium constant defines *quantitatively* the effect of concentrations of reactants and products on the extent of reaction. The manner of calculating such effects will be discussed later in the chapter. At present we may deduce qualitatively certain conclusions of general validity. For this purpose consider the reaction

$$H_2(g) + Cl_2(g) = 2\,HCl(g) \tag{12}$$
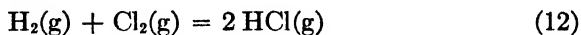
for which the equilibrium constant is given by

$$\frac{P^2_{HCl}}{P_{H_2}P_{Cl_2}} = K_p \tag{13}$$

Since K_p is a constant at all pressures, the relation among the partial pressures of H_2, Cl_2, and HCl in equation (13) should also remain constant under all conditions at a given temperature. If now hydrogen is added to an equilibrium mixture of the three gases, the pressure of this gas is increased, and hence the existing relationship among the partial pressures as given in equation (13) is disturbed. To accommodate the

added hydrogen without violating the constancy of K_p, the partial pressure of chlorine must decrease, while that of hydrogen chloride must increase. This can be accomplished by further interaction of hydrogen and chlorine to form hydrogen chloride, the process continuing until the relationship called for by K_p is reestablished and the gases are again in equilibrium. The same effect may be produced by adding chlorine. Conversely, addition of hydrogen chloride will increase the numerator, and hence the denominator must also be increased to preserve the constancy of K_p. This time the adjustment is brought about by dissociation of HCl into H_2 and Cl_2 until equilibrium is again reestablished and K_p returns to its constant value.

From this behavior of an equilibrium mixture on addition of excess reactants or products may be drawn the following two generalizations:

(1) Presence of excess of some of the reactants over others tends to drive a reaction further to completion with respect to the reactants not in excess.

(2) Initial presence of products tends to decrease the extent of conversion of reactants to products.

The Thermodynamic Approach to Equilibrium. The problem of chemical equilibrium and the equilibrium constant can be approached not only from the kinetic point of view through the law of mass action, but also thermodynamically. The latter approach is significant not only for the relations it establishes between free energy change and the equilibrium constants, but also for the light it sheds on the meaning of the term "active mass."

For any reaction such as

$$\alpha A + bB \cdots = cC + dD \cdots$$
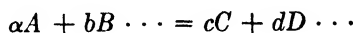
the change in free energy due to reaction, ΔF, at any temperature T is given by the reaction isotherm, equation (37) of the last chapter, namely,

$$\Delta F = \Delta F^0 + RT \ln \left(\frac{a_C^c a_D^d \cdots}{a_A^\alpha a_B^b \cdots} \right) \tag{14}$$

The activities indicated are those of the products at the end and of reactants at the start of the reaction. However, we have seen in the last chapter that the criterion of equilibrium is that $\Delta F = 0$. Hence, at equilibrium equation (14) becomes

$$0 = \Delta F^0 + RT \ln \left(\frac{a_C^c a_D^d \cdots}{a_A^\alpha a_B^b \cdots} \right)$$

$$\Delta F^0 = - RT \ln \left(\frac{a_C^c a_D^d \cdots}{a_A^\alpha a_B^b \cdots} \right) \tag{15}$$

The activities now are those of reactants and products *at equilibrium.* Since at any given temperature ΔF^0, the free energy change in the standard state, for any reaction is a constant, it follows that the activity ratio in equation (15) must also be a constant, or

$$\frac{a_C^c a_D^d \cdots}{a_A^a a_B^b \cdots} = K_a \tag{16}$$

and equation (15) may be written as

$$\Delta F^0 = - RT \ln K_a \tag{17}$$

K_a is the equilibrium constant of a reaction expressed in activities, and is known as the *thermodynamic equilibrium constant.* Equation (16), which is exact, indicates that only when activities are employed instead of concentrations or pressures can the equilibrium constant principle be expected to be generally valid. "Active mass," therefore, as far as equilibrium calculations are concerned, must be activity. Use of concentrations or pressures through equations (8) and (10) can represent only an approximation to equation (16), and will be equivalent to it only when activity is equal to concentration or pressure, i.e., at low pressures and concentrations. At higher pressures or concentrations K_p and K_c may exhibit nonconstancy due, not to the breakdown of the mass action principle, but to the fact that the pressure and concentration are no longer equal to the activity.

The limits within which concentrations and pressures may be substituted for activities have been discussed in the preceding chapter. Since no electrolytic equilibria will be discussed in this chapter, but only equilibria in gases and nonelectrolytic solutions, we shall find that concentrations and pressures in conjunction with equations (8) and (10) will suffice for our purposes, and will give results in essential agreement with equation (16).

Equation (17) relates directly the thermodynamic equilibrium constant of a reaction at a temperature T to the free energy change in the standard state for the reaction. It permits, therefore, the calculation of equilibrium constants from ΔF^0 values, and vice versa, the calculation of ΔF^0 of reactions from their equilibrium constants. It is this highly important equation which makes possible the reduction of all the free energy change calculations described in the last chapter to equilibrium constants, and allows prediction of the behavior of chemical reactions under various conditions without recourse to direct experiment. Thus, we have seen in the last chapter that for the reaction

$$\frac{3}{2} H_2(g) + \frac{1}{2} N_2(g) = NH_3(g) \qquad \Delta F^0_{298.2^\circ K} = - 3980 \text{ cal}$$
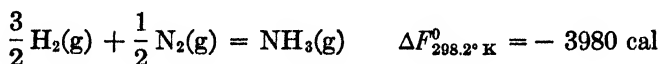
Using equation (17), we find, for the equilibrium constant of this reaction at 298.2° K,

$$\log_{10} K_a = -\frac{\Delta F^0}{2.303\ RT}$$

$$\frac{\cdot\ 3980}{2.303 \times 1.987 \times 298.2}$$

$$= 2.917$$

$$K_a = 826.1$$

Equilibria in Gaseous Systems. Many direct experimental studies of gaseous equilibria have been made by various investigators. Several examples of these will be discussed in detail to indicate some of the methods employed, as well as to illustrate the application of equilibrium constants.

The Ammonia Equilibrium. The equilibrium

$$\frac{3}{2}\,H_2(g) + \frac{1}{2}\,N_2(g) = NH_3(g) \tag{18}$$
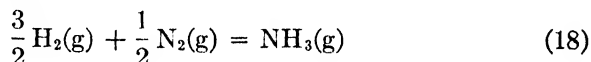
has been extensively investigated by Haber and Van Oordt, Haber and Rossighol, Nernst and Jellinck, and more recently by Larson and Dodge.[1] The last named passed a mixture of nitrogen and hydrogen, in the ratio of 1 : 3, through an iron coil immersed in a constant temperature bath. To speed the approach to equilibrium, the coil was lined with finely divided iron to act as a catalyst, i.e., a substance which accelerated the attainment of equilibrium without affecting it. The exit gases were analyzed, then, for hydrogen, nitrogen, and ammonia to determine the composition of the equilibrium mixture. To check their results and to assure themselves that true equilibrium has been attained, Larson and Dodge approached equilibrium also from the ammonia side by passing through the coil mixtures of ammonia, nitrogen, and hydrogen, allowing equilibrium to be established, and again analyzed the exit gases. Some of their results for the direct formation of ammonia from nitrogen and hydrogen at various pressures and temperatures are given in Table 1.

TABLE 1

AMMONIA EQUILIBRIUM CONSTANTS

$t°$ C	10 atm		30 atm		50 atm	
	% NH$_3$	K_p	% NH$_3$	K_p	% NH$_3$	K_p
350	7.35	0.0266	17.80	0.0273	25.11	0.0278
400	3.85	0.0129	10.09	0.0129	15.11	0.0130
450	2.04	0.00659	5.80	0.00676	9.17	0.00690

[1] Larson and Dodge, J. Am. Chem. Soc., **45**, 2918 (1923).

The first column indicates the temperature at which equilibrium was determined, column 2 gives the mol percentage of ammonia found in the equilibrium mixture under a total equilibrium pressure of ten atmospheres, while column 3 gives the value of the equilibrium constants K_p, as calculated from

$$K_p = \frac{P_{NH_3}}{P_{H_2}^{3/2} P_{N_2}^{1/2}} \tag{19}$$

The other columns give the same information under different total equilibrium pressures.

From the percentage of ammonia at equilibrium at a total pressure P, K_p is calculated with the aid of Dalton's law of partial pressures. Taking the data for 30 atmospheres and 400° C, we see that the percentage of ammonia at equilibrium is 10.09, and hence the partial pressure of ammonia is

$$P_{NH_3} = 30 \times 0.1009 = 3.03 \text{ atm}$$

The pressure of hydrogen plus nitrogen is, therefore,

$$P_{H_2} + P_{N_2} = 30.00 - 3.03 = 26.97 \text{ atm}$$

Since, however, the nitrogen and hydrogen are present in the ratio of 1 : 3, then,

$$P_{H_2} = \frac{3}{4} \times 26.97 = 20.22 \text{ atm}$$

$$P_{N_2} = \frac{1}{4} \times 26.97 = 6.75 \text{ atm}$$

and

$$\begin{aligned} K_p &= \frac{P_{NH_3}}{P_{H_2}^{3/2} P_{N_2}^{1/2}} \\ &= \frac{3.03}{(20.22)^{3/2}(6.75)^{1/2}} \\ &= 0.0129 \end{aligned}$$

Table 1 clearly substantiates the expression derived for the equilibrium constant on the basis of the law of mass action. Although the variation in total pressure is quite large, 10 to 50 atmospheres, still K_p at any *given temperature* is essentially constant. The total pressure merely determines the relative percentages of hydrogen, nitrogen, and ammonia present at equilibrium, but does not in any way affect the constancy of K_p. On the other hand, variation of temperature produces significant changes both in the percentage of ammonia and in K_p. At all the given pressures increase of temperature operates to decrease the yield of ammonia and hence to decrease K_p.

The Le Chatelier-Braun Principle. To foretell qualitatively the effect of variation in pressure or temperature on a system in equilibrium,

use is made of the *Le Chatelier-Braun principle*. This principle states that *whenever stress is placed on any system in a state of equilibrium, the system will always react in a direction which will tend to counteract the applied stress*. Thus, if pressure is applied to a system, the tendency of the stress will be to decrease the volume, and hence that reaction in the system will take place which will favor the smaller volume. In the ammonia equilibrium the combination of nitrogen and hydrogen to form ammonia is attended by a volume *decrease* from two molar volumes to one molar volume, and hence according to the Le Chatelier-Braun principle we may expect the formation of ammonia to be favored by an increase in total pressure. That this is actually the case is borne out by Table 1.

Again, when a reaction is endothermic, i.e., absorbs heat, addition of heat should favor it, and the reaction should proceed to a greater extent at higher temperatures. Alternately, when a reaction is exothermic we may expect that addition of heat would tend to inhibit the reaction, and hence at higher temperatures the reaction should tend to reverse itself. The latter behavior is shown by the ammonia synthesis reaction, which is exothermic. Table 1 indicates, in line with the Le Chatelier-Braun prediction, that the yield of ammonia is higher the lower the temperature.

The Phosgene Equilibrium. The formation of phosgene,

$$CO(g) + Cl_2(g) = COCl_2(g) \tag{20}$$
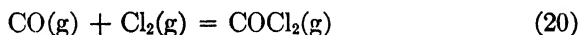
has been fully studied by Max Bodenstein and Heinrich Plaut [1] by a static method, as contrasted to the flow, or dynamic, method employed by Larson and Dodge for the ammonia equilibrium. The apparatus used consisted of a glass reaction bulb set in a cylindrical electrically heated oven which was kept at a constant temperature. The top of the glass bulb was connected by capillary glass tubing to tanks in which the various gases were stored, and to a specially constructed manometer made of quartz. All pressure measurements were made with this manometer. At the start of an experiment chlorine was first admitted, then carbon monoxide, and the pressure of each was recorded. After equilibrium had been established, the total pressure as recorded on the manometer was also read.

K_p can be calculated directly from the pressure readings. In a typical experiment Bodenstein and Plaut found at 394.8° C that for an initial pressure of chlorine equal to 351.4 mm and carbon monoxide equal to 342.0 mm, the total pressure at equilibrium was 439.5 mm. Since the volume was constant throughout the experiment, the partial pressures are directly proportional to the numbers of moles of each constituent present, and hence we may deal with these directly. If, then, we let x

[1] Bodenstein and Plaut, Z. physik. Chem., **110**, 399 (1924).

be the drop in the partial pressure of chlorine during the experiment, the drop in partial pressure of carbon monoxide is also x, while the partial pressure of phosgene formed is x. This relation is apparent from the stoichiometry of equation (20), where the two gases, chlorine and carbon monoxide, interact mole for mole to form 1 mole of phosgene. We may write, therefore, for the partial pressures of the three gases at equilibrium:

$$P_{Cl_2} = 351.4 - x \text{ mm}$$
$$P_{CO} = 342.0 - x \text{ mm}$$
$$P_{COCl_2} = x \text{ mm}$$

The total pressure at equilibrium must be the sum of these partial pressures, and must equal in turn to 439.5 mm as found by experiment.

Then
$$\begin{aligned} P = P_{Cl_2} + P_{CO} + P_{COCl_2} &= 439.5 \text{ mm} \\ = (351.4 - x) + (342.0 - x) + x &= 439.5 \text{ mm} \\ = 693.4 - x &= 439.5 \text{ mm} \\ x = 693.4 - 439.5 &= 253.9 \text{ mm} \end{aligned}$$

Substituting this value of x into the expressions for the partial pressures, we find
$$\begin{aligned} P_{Cl_2} = 351.4 - 253.9 \\ = 97.5 \text{ mm} = 0.128 \text{ atm} \\ P_{CO} = 342.0 - 253.9 \\ = 88.1 \text{ mm} = 0.116 \text{ atm} \\ P_{COCl_2} = 253.9 \text{ mm} = 0.334 \text{ atm} \end{aligned}$$

Hence at 394.8° C

$$K_p = \frac{P_{COCl_2}}{P_{Cl_2}P_{CO}} = \frac{0.334}{(0.128)(0.116)}$$
$$= 22.5$$

When the equilibrium constant for a forward reaction is known, the equilibrium constant for the reverse reaction is also known through equation (9). Thus, since the equilibrium constant for the formation of phosgene from chlorine and carbon monoxide is $K_p = 22.5$ at 394.8° C, the equilibrium constant for the *dissociation* of phosgene into carbon monoxide and chlorine, namely,

$$COCl_2(g) = CO(g) + Cl_2(g) \tag{21}$$
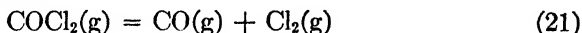
must be
$$K_p' = \frac{P_{CO}P_{Cl_2}}{P_{COCl_2}} = \frac{1}{K_p} \tag{22}$$
$$\frac{1}{22.5} = 0.0444$$

This new constant may be employed to calculate the extent to which phosgene dissociates into the given products at any specified pressure at 394.8° C.

If we start with n moles of phosgene, and let α be the degree of dissociation at equilibrium, i.e., the *fraction of each mole* that dissociates, the number of moles of phosgene undissociated at equilibrium must be $n(1 - \alpha)$, while the number of moles of chlorine and carbon monoxide formed must each be $n\alpha$. Hence the total number of moles of gas present at equilibrium is

$$n_t = n(1 - \alpha) + n\alpha + n\alpha$$
$$= (n + n\alpha)$$
$$= n(1 + \alpha)$$

If the total equilibrium pressure is P, the partial pressures of the three gases, using Dalton's law, must be

$$P_{COCl_2} = \left(\frac{n_{COCl_2}}{n_t}\right)P = \left(\frac{1 - \alpha}{1 + \alpha}\right)P$$

$$P_{Cl_2} = \left(\frac{n_{Cl_2}}{n_t}\right)P = \left(\frac{\alpha}{1 + \alpha}\right)P$$

$$P_{CO} = \left(\frac{n_{CO}}{n_t}\right)P = \left(\frac{\alpha}{1 + \alpha}\right)P$$

and therefore,

$$K_p' = \frac{P_{CO}P_{Cl_2}}{P_{COCl_2}} = \frac{\left(\dfrac{\alpha}{1 + \alpha}\right)P \cdot \left(\dfrac{\alpha}{1 + \alpha}\right)P}{\left(\dfrac{1 - \alpha}{1 + \alpha}\right)P}$$

$$= \frac{\alpha^2 P}{(1 - \alpha)(1 + \alpha)}$$

$$= \frac{\alpha^2 P}{(1 - \alpha^2)} \tag{23}$$

Substituting $K_p' = 0.0444$, and assuming $P = 1$ atmosphere, solution for α yields

$$\frac{\alpha^2}{1 - \alpha^2} = 0.0444$$
$$\alpha = 0.206$$

Consequently, pure phosgene dissociates into chlorine and carbon monoxide to the extent of 20.6 per cent at 1 atmosphere pressure and a temperature of 394.8° C.

This typical calculation illustrates the general procedure followed in setting up expressions for the equilibrium constant in terms of the total pressure and the degree of dissociation.

The Dissociation of Antimony Pentachloride. In the phosgene equilibrium the α's were ascertained from pressure measurements. In dissociations the α's may be deduced also from density or apparent molec-

ular weight data. To illustrate this method of calculation, consider the equilibrium

$$SbCl_5(g) = SbCl_3(g) + Cl_2(g) \tag{24}$$
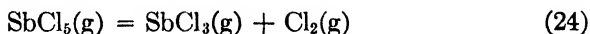
If we start with n moles of antimony pentachloride in a volume of V liters, and let α be the degree of dissociation at temperature T and total equilibrium pressure P, then the number of moles of antimony penta-chloride left undissociated is $n(1 - \alpha)$, the number of moles each of antimony trichloride and chlorine formed is $n\alpha$, and the total number of moles of gas present at equilibrium is thus $n_t = n(1 + \alpha)$. For such an equilibrium gas mixture, the ideal gas law gives the relation

$$\begin{aligned} PV &= n_t RT \\ &= n(1 + \alpha)RT \end{aligned} \tag{25}$$

Since n is the weight W of antimony pentachloride started with divided by the molecular weight M of this substance, and since the total mass of gas does not change, $W = W_{\text{mixture}}$, $n = W/M = W_{\text{mixture}}/M$, and equation (25) becomes

$$PV = \frac{W_{\text{mixture}}}{M}(1 + \alpha)RT$$

or

$$P = \left(\frac{W_{\text{mixture}}}{V}\right)\frac{(1 + \alpha)RT}{M}$$

$$= \frac{d_a(1 + \alpha)RT}{M} \tag{26}$$

where d_a is the density of the mixture. Solving for α, we find that

$$\alpha = \frac{PM}{d_a RT} - 1 \tag{27}$$

and hence α can be calculated from the observed pressure, temperature, and density of the mixture. It should be remembered that in equation (27) M is the normal molecular weight of antimony pentachloride as obtained from atomic weights. Furthermore, reference to equation (58), Chapter I, shows that PM/RT is what the density d of antimony pentachloride would have been at the given temperature and pressure if no dissociation had occurred. Substituting then d for PM/RT in equation (27), we obtain also for α

$$\alpha = \frac{d}{d_a} - 1$$

$$\alpha = \frac{d - d_a}{d_a} \tag{28}$$

Equation (28) may be cast into still another form. If the numerator and denominator are multiplied by the molar volume of a gas at the given temperature and pressure, dV_m becomes M, the molecular weight of the antimony pentachloride, d_aV_m becomes the apparent molecular weight M_a of the gas mixture, and equation (28) is then

$$\alpha = \frac{M - M_a}{M_a} \tag{29}$$

The apparent molecular weight of the equilibrium gas mixture is what would be calculated for the molecular weight of the antimony pentachloride if it were assumed that only n moles of gas were present at the temperature, pressure, and volume of the equilibrium gas mixture, i.e.,

$$PV = nRT$$
$$= \frac{WRT}{M_a} \tag{30}$$

In Table 2 are given some data on the dissociation of antimony pentachloride at various temperatures and 1 atmosphere pressure. Column 1 lists the temperature and column 2 the apparent molecular weight calculated from the observed data by means of equation (30), while in column 3 are the values of α obtained through equation (29). The molecular weight of antimony pentachloride is 299.05. In the interval 128 to 248° C, therefore, α increases from 8.2 to 71.8 per cent.

TABLE 2

Degree of Dissociation of SbCl₅ at Various Temperatures [1]

(Total pressure = 1 atm)

Temperature ° C	M_a	$\alpha = \dfrac{M - M_a}{M_a}$
128	276.2	0.082
141	265.5	0.126
157	252.7	0.183
169	245.3	0.219
182	231.5	0.292
191	221.3	0.351
206	204.4	0.462
221	194.2	0.540
233.5	179.3	0.668
248	174.0	0.718

Equations (28) and (29) derived for the dissociation of antimony pentachloride apply equally well to any other dissociations in which one molecule decomposes to yield *two* molecules of product. More generally

[1] Calculated from the data of Braune and Tiedje, Zeit. anorg. allgem. Chem., **152,** 39 (1926).

it can be shown that when a single molecule dissociated into ν moles of product, whether identical or not, equation (28) becomes

$$\alpha = \frac{d - d_a}{(\nu - 1)d_a} \tag{31}$$

and equation (29) becomes

$$\alpha = \frac{M - M_a}{(\nu - 1)M_a} \tag{32}$$

The Dissociation of Hydrogen Sulfide. Hydrogen sulfide on being heated dissociates into hydrogen and sulfur according to the equation

$$2\,H_2S(g) = 2\,H_2(g) + S_2(g) \tag{33}$$
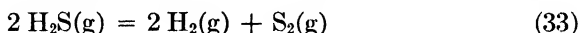
To determine the extent of this dissociation, Preuner and Schupp [1] utilized a very novel method. Hydrogen sulfide under a definite pressure was admitted into an elongated porcelain tube where it was allowed to dissociate at constant temperature. In the center of this tube was located a small platinum bulb which acted as a membrane permeable to hydrogen, but not to hydrogen sulfide or sulfur. This bulb was evacuated before each experiment. As the hydrogen sulfide dissociated, the hydrogen formed diffused rapidly into the platinum bulb until it built up within a pressure equal to the equilibrium pressure of hydrogen. This pressure was recorded on a manometer connected directly with the bulb, and was used to calculate the degree of dissociation, α, of the hydrogen sulfide. Values of α obtained in this manner at various temperatures for a total equilibrium pressure of *one* atmosphere are shown in Table 3.

TABLE 3

DISSOCIATION OF H_2S AT 1 ATMOSPHERE PRESSURE

$t°$ C	α	K_p
750	0.055	0.000091
830	0.087	0.00038
1065	0.247	0.0118
1132	0.307	0.0260

To calculate K_p from these values of α we proceed as follows. Of each initial mole of hydrogen sulfide the amount left at equilibrium is $(1 - \alpha)$, and of n moles, $n(1 - \alpha)$. Since for each mole of hydrogen sulfide that dissociates 1 mole of hydrogen and 0.5 mole of sulfur are formed, then from a total of $n\alpha$ moles dissociated, $n\alpha$ moles of hydrogen and $n\alpha/2$ moles of sulfur are obtained. The conditions at equilibrium are, therefore,

$$2\,H_2S \;=\; 2\,H_2 + S_2$$
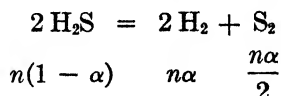
$$n(1 - \alpha) \qquad n\alpha \qquad \frac{n\alpha}{2}$$

[1] Preuner and Schupp, Z. physik. Chem., **68**, 157 (1909).

and instead of the initial n moles of gas there are present at equilibrium

$$n_t = n(1 - \alpha) + n\alpha + \frac{n\alpha}{2}$$

$$= n + \frac{n\alpha}{2}$$

$$= \frac{n(2 + \alpha)}{2}$$

For a total equilibrium pressure P, the partial pressures are, respectively,

$$P_{\mathrm{H_2S}} = \left[\frac{n(1 - \alpha)}{n\dfrac{(2 + \alpha)}{2}}\right] P = \left[\frac{2(1 - \alpha)}{2 + \alpha}\right] P$$

$$P_{\mathrm{H_2}} = \left[\frac{n\alpha}{n\dfrac{(2 + \alpha)}{2}}\right] P = \left[\frac{2\,\alpha}{2 + \alpha}\right] P$$

$$P_{\mathrm{S_2}} = \left[\frac{\left(\dfrac{n\alpha}{2}\right)}{n\left(\dfrac{2 + \alpha}{2}\right)}\right] P = \left[\frac{\alpha}{2 + \alpha}\right] P$$

and, therefore,
$$K_p = \frac{P_{\mathrm{H_2}}^2 P_{\mathrm{S_2}}}{P_{\mathrm{H_2S}}^2} = \frac{\left[\dfrac{2\,\alpha}{2 + \alpha}\right]^2 P^2 \cdot \left[\dfrac{\alpha}{2 + \alpha}\right] P}{\left[\dfrac{2(1 - \alpha)}{(2 + \alpha)}\right]^2 P^2}$$

$$= \frac{\alpha^3 P}{(2 + \alpha)(1 - \alpha)^2} \tag{34}$$

The values of K_p obtained by substituting $P = 1$ and the α's from Table 3 into the above expression are given in column 3 of the table.

Expressions for K_p like equations (23) or (34) apply only to the equilibrium conditions for which they were derived. When the conditions under which equilibrium is established are varied, such as when products as well as reactants are present initially, the expressions are different in form and more complicated. Thus, consider the hydrogen sulfide equilibrium when $n_{\mathrm{H_2S}}$ moles of hydrogen sulfide, $n_{\mathrm{H_2}}$ moles of hydrogen, and $n_{\mathrm{S_2}}$ moles of sulfur are mixed at 1132° C. To obtain the equilibrium relations at a total pressure P, let x be the number of moles of hydrogen

sulfide which react to form hydrogen and sulfur. The number of moles of each specie at equilibrium is, then,

$$2\,H_2S \;=\; 2\,H_2 \;+\; S_2$$
$$(n_{H_2S} - x) \quad (n_{H_2} + x) \quad \left(n_{S_2} + \frac{x}{2}\right)$$

and

$$n_t = (n_{H_2S} - x) + (n_{H_2} + x) + \left(n_{S_2} + \frac{x}{2}\right)$$
$$= \left(n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}\right)$$

Consequently, the partial pressures are

$$P_{H_2S} = \left(\frac{n_{H_2S} - x}{n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}}\right) P$$

$$P_{H_2} = \left(\frac{n_{H_2} + x}{n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}}\right) P$$

$$P_{S_2} = \left(\frac{n_{S_2} + \frac{x}{2}}{n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}}\right) P$$

and the expression for K_p becomes

$$K_p = \frac{P^2_{H_2} P_{S_2}}{P^2_{H_2S}}$$

$$= \frac{\left(\dfrac{n_{H_2} + x}{n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}}\right)^2 P^2 \cdot \left(\dfrac{n_{S_2} + \frac{x}{2}}{n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}}\right) P}{\left(\dfrac{n_{H_2S} - x}{n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}}\right)^2 P^2}$$

$$= \frac{(n_{H_2} + x)^2 \left(n_{S_2} + \frac{x}{2}\right) P}{\left(n_{H_2S} + n_{H_2} + n_{S_2} + \frac{x}{2}\right)(n_{H_2S} - x)^2} \tag{35}$$

For the *special* case of $n_{H_2S} = n_{H_2} = n_{S_2} = 1$, and $P = 1$ atmosphere, the equilibrium constant expression at $1132°$ C, where $K_p = 0.0260$, reduces to

$$K_p = \frac{(1+x)^2\left(1+\dfrac{x}{2}\right)}{\left(3+\dfrac{x}{2}\right)(1-x)^2} = 0.0260$$

Solution of this equation for x yields $x = -0.526$. The minus sign indicates that hydrogen sulfide does *not* dissociate under the conditions specified, but that hydrogen and sulfur must combine to form hydrogen sulfide before equilibrium can be established in a mixture containing initially 1 mole of each of the participants and at a total equilibrium pressure of 1 atmosphere. At equilibrium there are present, therefore,

$$n_{H_2S} - x = 1 + 0.526 = 1.526 \text{ moles of } H_2S$$
$$n_{H_2} + x = 1 - 0.526 = 0.474 \text{ mole of } H_2$$
$$n_{S_2} + \frac{x}{2} = 1 - 0.263 = 0.737 \text{ mole of } S_2$$

and the partial pressures are, respectively,

$$P_{H_2S} = \frac{1.526}{2.737} \times 1 = 0.558 \text{ atm}$$
$$P_{H_2} = \frac{0.474}{2.737} \times 1 = 0.173 \text{ atm}$$
$$P_{S_2} = \frac{0.737}{2.737} \times 1 = 0.269 \text{ atm}$$

Effect of Inert Gases on Equilibrium. All the examples of equilibria considered so far involved only the gases which participate directly in the reaction. Frequently, however, equilibrium mixtures are encountered where other gases are present as well as those involved in the reaction. The question is: What, if any, will be the effect of gases not entering into the reaction on the extent of reaction? Obviously, the presence of these gases cannot affect in any way the equilibrium constant; but it can and does change the *partial pressures* of reactants and products, and hence we may expect a shift in the extent of reaction to permit a redistribution of the partial pressures in accord with the demands of the equilibrium constant.

To illustrate quantitatively the effect of an inert gas on an equilibrium, consider again the reaction

$$COCl_2(g) = CO(g) + Cl_2(g)$$

We have seen that at 394.8° C and at a total equilibrium pressure of 1 atmosphere the degree of dissociation of the phosgene is $\alpha = 0.206$. Suppose now, however, that equilibrium is established in presence of nitrogen gas at a partial pressure of 0.40 atmosphere in a total pressure of 1 atmosphere. Then the sum of the partial pressures of phosgene, chlorine, and carbon monoxide is no longer $P = 1$, but $1 - 0.40 = 0.60$ atmosphere, and from equation (23) it follows that

$$K_p = \frac{P_{CO} \times P_{Cl_2}}{P_{COCl_2}} = \frac{\alpha^2 P}{(1 - \alpha^2)}$$

$$0.0444 = \frac{\alpha^2(0.60)}{(1 - \alpha^2)}$$

$$\alpha = 0.262$$

The addition of nitrogen under the specified conditions leads, therefore, to an increase of 5.6 per cent in the dissociation of phosgene.

Qualitatively, the effect of inert gases on the extent of reaction can be foretold with the aid of the Le Chatelier-Braun principle. Since for any given total pressure the presence of inert gases decreases the partial pressures of reactants and products, the net effect will be the same as if the gases at equilibrium were subjected to a lower total pressure. Or, stated differently, the effect of an inert gas is to *dilute* the concentrations of reactants and products. Consequently, according to the Le Chatelier-Braun principle, presence of inert gases will favor the reaction that results in an *increase* in volume. Thus, when the volume of reactants is greater than that of products, the reaction will be displaced in favor of reactants, while when the opposite is true, the formation of products will be favored, as is substantiated by the above calculation. When there is no change in volume on reaction, however, neither side will be favored, and the equilibrium will be uninfluenced by the introduction of inert gases.

Equilibria in Liquid Systems. As an example of an equilibrium in a liquid solution may be taken the dissociation of the amyl ester of dichloracetic acid into the acid and amylene, namely,

$$CHCl_2COOC_5H_{11} = CHCl_2COOH + C_5H_{10} \tag{37}$$

for which the equilibrium constant K_c is given by

$$K_c = \frac{(C_{acid})(C_{amylene})}{C_{ester}} \tag{38}$$

In investigating this equilibrium Nernst and Hohmann[1] mixed various proportions of amylene and dichloracetic acid, sealed the mixtures in

---

[1] Nernst and Hohmann, Z. physik. Chem., **11**, 352 (1893).

glass tubes, and kept the tubes at 100° C until equilibrium was established. The tubes were chilled then to "freeze" the equilibrium, opened, and the contents analyzed for the amount of ester present.

If we let a equal the initial number of moles of acid, b the initial number of moles of amylene, x the number of moles of ester at equilibrium, and V the total volume of mixture in liters, then the concentrations of the various substances at equilibrium in moles per liter are

$$C_{ester} = \frac{x}{V}$$

$$C_{acid} = \frac{a - x}{V}$$

$$C_{amylene} = \frac{b - x}{V}$$

and, therefore,

$$K_c = \frac{\left(\dfrac{a - x}{V}\right)\left(\dfrac{b - x}{V}\right)}{\left(\dfrac{x}{V}\right)}$$

$$= \frac{(a - x)(b - x)}{xV} \tag{39}$$

In all of Nernst and Hohmann's experiments a was kept at 1 mole, while b was varied. The values b, V, and x for a series of runs are shown in the first three columns of Table 4, while the fourth column shows the values of K_c as calculated from these data by equation (39). The constancy in K_c is fairly satisfactory.

TABLE 4

DECOMPOSITION OF THE AMYL ESTER OF DICHLORACETIC ACID AT 100° C

(a = 1 mole)

b (moles)	V (liters)	x	K_c
1.05	0.215	0.455	3.31
2.61	0.401	0.615	3.12
4.45	0.640	0.628	3.54
5.91	0.794	0.658	3.44
7.30	0.959	0.650	3.73
8.16	1.062	0.669	3.49
11.33	1.439	0.688	3.35
13.80	1.734	0.700	3.24
15.36	1.829	0.703	3.39

Often the values of K_c obtained in the study of equilibria in solutions are not particularly constant, but deviate more and more as the concen-

trations of the various constituents in solution are increased. The deviations usually are not very serious with nonelectrolytes or weak electrolytes, but they are very pronounced when strong electrolytes are involved. In such cases true equilibrium constants can be obtained only by the use of activities in place of concentrations. Homogeneous equilibria involving strong electrolytes will be discussed in a subsequent chapter.

Equilibrium Constants and Temperature. Although the equilibrium constant of a reaction is constant at any given temperature, its magnitude varies appreciably as the temperature is changed. The extent of this variation for a typical reaction may be judged from Table 3. The exact manner in which the equilibrium constant of any reaction depends on the temperature is controlled by the thermodynamic relations of the reaction and is readily deducible from these.

If equation (17) of this chapter is differentiated with respect to temperature, then

$$\Delta F^0 = - RT \ln K_a$$

$$\left[\frac{\partial \Delta F^0}{\partial T} \right] = - RT \left[\frac{\partial \ln K_a}{\partial T} \right] - R \ln K_a$$

But, according to the Gibbs-Helmholtz equation,

$$\left[\frac{\partial \Delta F^0}{\partial T} \right] = \frac{\Delta F^0 - \Delta H}{T} = \frac{- RT \ln K_a - \Delta H}{T}$$

$$= - R \ln K_a - \frac{\Delta H}{T}$$

Therefore,

$$- \cancel{R \ln K_a} - \frac{\Delta H}{T} = - RT \left[\frac{\partial \ln K_a}{\partial T} \right] - \cancel{R \ln K_a}$$

and

$$\left[\frac{\partial \ln K_a}{\partial T} \right] = \frac{\Delta H}{RT^2} \tag{40}$$

Equation (40), known as the *van't Hoff reaction isobar*, defines the temperature coefficient of $\ln K_a$ in terms of the heat of reaction ΔH and the temperature T. For gaseous reactions, when $K_a = K_p$, equation (40) may be written as

$$\left[\frac{\partial \ln K_p}{\partial T} \right] = \frac{\Delta H}{RT^2} \tag{41}$$

To obtain a solution for $\ln K_p$ from equation (41) the equation must be integrated. For exact integration ΔH must be known as a function of T. However, when the temperature interval considered is not very

large, ΔH may be considered constant over the interval, and

$$\int_{K_{p_1}}^{K_{p_2}} d \ln K_p = \int_{T_1}^{T_2} \frac{\Delta H}{RT^2} dT$$

$$\ln K_p \Big]_{K_{p_1}}^{K_{p_2}} = \frac{\Delta H}{R} \left[-\frac{1}{T} \right]_{T_1}^{T_2}$$

$$\ln \frac{K_{p_2}}{K_{p_1}} = \frac{\Delta H}{R} \left[\frac{T_2 - T_1}{T_1 T_2} \right] \qquad (42)$$

Equation (42) permits the calculation of K_{p_2} at T_2 when K_{p_1} at T_1 and ΔH are available; or, when the equilibrium constants at two different temperatures are known, equation (42) may be used to obtain the average heat of the reaction over the temperature range T_1 to T_2. These calculations may be illustrated by the following two examples.

(1) The equilibrium constant for the dissociation

$$2 H_2S(g) = 2 H_2(g) + S_2(g)$$
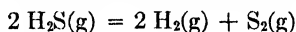
is $K_p = 0.0118$ at $1065°$ C, while the heat of dissociation is $\Delta H = 42,400$ calories. Find the equilibrium constant of the reaction at $1200°$ C. Applying equation (42),

$$\log_{10} \frac{K_{p_2}}{K_{p_1}} = \frac{\Delta H}{2.303 \times R} \left[\frac{T_2 - T_1}{T_1 T_2} \right]$$

$$\log_{10} \frac{K_{p_2}}{0.0118} = \frac{42,400}{4.57} \left[\frac{1473 - 1338}{1473 \times 1338} \right]$$

$$= 0.6350$$

$$\frac{K_{p_2}}{0.0118} = 4.30$$

$$K_{p_2} = 4.30 \times 0.0118$$

$$= 0.0507 \text{ at } 1200° \text{ C}$$

(2) For the reaction

$$SO_2(g) + \frac{1}{2} O_2(g) = SO_3(g)$$
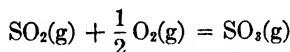
$K_p = 6.55$ at $900°$ K and $K_p = 1.86$ at $1000°$ K. Calculate the heat of the reaction over the temperature interval 900 to 1000° K.

Again applying equation (42), we find

$$\log_{10} \frac{1.86}{6.55} = \frac{\Delta H}{4.57} \left[\frac{1000 - 900}{900 \times 1000} \right]$$

$$\Delta H = -22,500 \text{ cal}$$

Instead of integrating equation (41) between limits, a general solution under the same conditions may be obtained in the form

$$d \ln K_p = \frac{\Delta H}{RT^2} dT$$

$$\ln K_p = -\frac{\Delta H}{RT} + C \qquad (43)$$

C is an integration constant, and may be evaluated for any reaction by substituting a known value of K_p at some given temperature. When several values of K_p at various temperatures are available and ΔH is sought, it is preferable to determine ΔH graphically through equation (43) rather than to use equation (42). According to equation (43) a plot of $\log_{10} K_p$ vs. $1/T$ should be a straight line with slope equal to $-\Delta H/2.303\,R$, and hence ΔH follows from such a plot as

$$\Delta H = -2.303\,R \times \text{slope}$$
$$= -4.57 \times \text{slope} \tag{44}$$

Table 5 lists the equilibrium constant at various temperatures for the reaction

$$SO_2(g) + \frac{1}{2} O_2(g) = SO_3(g) \tag{45}$$
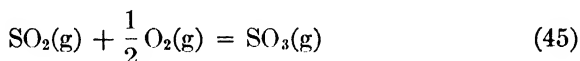
while Fig. 1 gives a plot of $\log_{10} K_p$ vs. $1/T$ constructed from these data. In agreement with equation (43) the plot is a straight line with slope equal to 4930. Consequently,

$$\Delta H = -4.57 \times 4930$$
$$= -22,500 \text{ cal}$$

TABLE 5

EQUILIBRIUM CONSTANTS AT VARIOUS TEMPERATURES
FOR THE REACTION: $SO_2(g) + 1/2\,O_2(g) = SO_3(g)$

T° K	K_p	$1/T$	$\log_{10} K_p$
801	31.3	0.001248	1.496
852	13.8	0.001174	1.140
900	6.55	0.001111	0.816
953	3.24	0.001049	0.511
1000	1.86	0.001000	0.270
1062	0.96	0.000942	−0.018
1105	0.63	0.000905	−0.202
1170	0.36	0.000855	−0.444

K_c and Temperature. An equation similar to (41) may also be derived for the variation of K_c of a gaseous reaction with temperature, namely,

$$\left[\frac{\partial \ln K_c}{\partial T}\right] = \frac{\Delta E}{RT^2} \tag{46}$$

where ΔE is the heat of reaction at constant volume. This equation is integrated and handled in exactly the same manner as equation (41). As a general rule equations (46) and (41) will yield different results, and will be identical only when $K_c = K_p$ and $\Delta H = \Delta E$, i.e., when $\Delta n = 0$ and there is no volume change on reaction.

Equation (46) is valid for reactions in solution. For such reactions the difference between ΔH and ΔE is negligible, and hence equation (46) may be written also

$$\left[\frac{\partial \ln K_c}{\partial T}\right] = \frac{\Delta E}{RT^2} = \frac{\Delta H}{RT^2} \tag{47}$$

The Nernst Approximation Formula. When neither equilibrium nor free energy data are available to permit the exact calculation of the equilibrium constant of a reaction, a rough *estimate* of the equilibrium
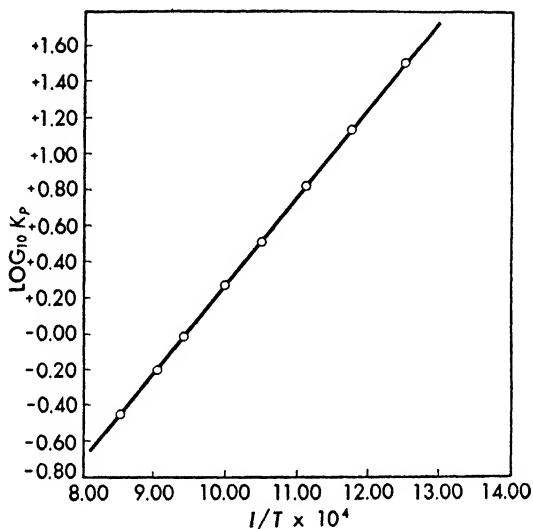
Fig. 1. Plot of $\text{Log}_{10} Kp$ vs. $1/T$ for the Reaction:
$$SO_2(g) + \tfrac{1}{2} O_2(g) = SO_3(g)$$

constant can still be made provided the heat of the reaction is known. This is made possible by the use of an empirical equation suggested by Walter Nernst, and called the *Nernst approximation formula*, namely,

$$\log_{10} K_p = -\frac{\Delta H}{4.57\,T} + 1.75\,\Delta n\,\log_{10} T + \Sigma(nC) \tag{48}$$

In this equation K_p is the equilibrium constant, ΔH the heat of reaction, T the absolute temperature, and Δn the difference between the number of moles of *gaseous* products and the number of moles of *gaseous* reactants involved in the reaction. C is an empirical constant characteristic for each *gaseous* substance, and is known as the *conventional chemical constant* of the substance, while n represents the number of moles of each *gaseous* specie participating in the reaction. The term $\Sigma(nC)$ is defined as

$$\Sigma(nC) = (n_1C_1 + n_2C_2 \cdots)_{\text{products}} - (n_1'C_1' + n_2'C_2' \cdots)_{\text{reactants}} \tag{49}$$

This formula may be applied to reactions involving gases only, as well as to reactions in which solids and liquids take part, *provided the presence of the solid and liquid phases is completely disregarded in evaluating Δn and $\Sigma(nC)$.* Only the gaseous phases need be considered. The conventional chemical constants required are given in Table 6 for a number of substances. It will be observed that the constants of all substances except hydrogen are very close to 3.0; hence, 3.0 may be assumed for the chemical constant of any substance for which no constant is available.

TABLE 6

Conventional Chemical Constants

H_2	1.6	HCl	3.0	N_2O	3.3	CH_4	2.8
N_2	2.6	HBr	3.2	NH_3	3.3	C_2H_6	2.6
O_2	2.8	HI	3.4	H_2O	3.6	C_2H_4	2.8
Cl_2	3.1	CO	3.5	H_2S	3.0	C_2H_2	3.2
I_2	3.9	CO_2	3.2	CS_2	3.1	C_6H_6	3.0
Br_2	3.2	NO	3.5	SO_2	3.3	CCl_4	3.4

Example: Using the Nernst approximation formula, find the equilibrium constant at 291° K for the reaction

$$CO_2(g) + C(s) = 2\,CO(g) \qquad \Delta H_{291} = 41{,}950 \text{ cal} \qquad (50)$$

For this reaction, $\Delta n = 2 - 1 = 1$, while $\Sigma(nC)$, using the constants in Table 6, is

$$\begin{aligned} \Sigma(nC) &= 2\,C_{CO} - C_{CO_2} \\ &= 2 \times 3.5 - 3.2 \\ &= 3.8 \end{aligned}$$

Therefore,

$$\begin{aligned} \log_{10} K_p &= \frac{-\Delta H}{4.57\,T} + 1.75\,\Delta n \log_{10} T + \Sigma(nC) \\ &= \frac{-41{,}950}{4.57 \times 291} + 1.75 \log_{10} 291 + 3.8 \\ &= -23.38 \end{aligned}$$

It is obvious that K_p is very small, and hence the reaction hardly takes place at room temperature.

REFERENCES FOR FURTHER READING

1. J. A. V. Butler, "Fundamentals of Chemical Thermodynamics," The Macmillan Company, New York, 1934–1935, 2 vol.
2. Hougen and Watson, "Industrial Chemical Calculations," second edition, John Wiley & Sons, Inc., New York, 1936.
3. K. Jellinek, "Lehrbuch der physikalischen Chemie," Ferdinand Enke, Stuttgart, 1930, Vol. III.
4. Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.

5. F. H. MacDougall, "Thermodynamics and Chemistry," third edition, John Wiley & Sons, Inc., New York, 1939.

6. Parks and Huffman, "The Free Energies of Organic Compounds," Reinhold Publishing Corporation, New York, 1932.

7. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931.

PROBLEMS

1. State which of the following equilibria are homogeneous and which are heterogeneous:

    (a) S (rhombic) = S (monoclinic)

    (b) $Fe_2O_3(s) + CO(g) = CO_2(g) + 2 FeO(s)$

    (c) $2 SO_2(g) + O_2(g) = 2 SO_3(g)$

    (d) $CaCO_3(s) + H^+(aq) = HCO_3^-(aq) + Ca^{++}(aq)$

2. Formulate the equilibrium constant K_p for each of the following reactions:

    (a) $C_2H_6(g) = C_2H_4(g) + H_2(g)$

    (b) $2 NO(g) + O_2(g) = 2 NO_2(g)$

    (c) $NO_2(g) + SO_2(g) = SO_3(g) + NO(g)$

    (d) $3 O_2(g) = 2 O_3(g)$

3. In each of the reactions of the preceding problem determine the ratio of K_p to K_c at 27° C.

4. The standard free energy of formation of HCl at 25° C is − 22,700 cal/mole. Calculate the thermodynamic equilibrium constant for the dissociation of HCl into its elements at 25° C.   *Ans.* 2.28×10^{-17}

5. In each of the following equilibria, predict qualitatively the effect of increasing the total pressure upon the percentage of products present at equilibrium:

    (a) $2 SO_3(g) + heat = 2 SO_2(g) + O_2(g)$

    (b) $2 HI(g) = H_2(g) + I_2(g) + heat$

    (c) $2 NO_2(g) = N_2O_4(g) + heat$

    (d) $CO(g) + H_2O(g) = CO_2(g) + H_2(g) + heat$

6. In the preceding problem, what effect will a decrease in temperature have upon the percentage of product present at equilibrium?

7. At 2155° C and 1 atm pressure steam is 1.18% decomposed into H_2 and O_2, according to the equation: $2 H_2O(g) = 2 H_2(g) + O_2(g)$. Calculate K_p for the decomposition of steam at this temperature.   *Ans.* 8.38×10^{-7}

8. At a total pressure of 1 atm and at 184° C, NO_2 is 5% decomposed according to the equation:

$$2 NO_2(g) = 2 NO(g) + O_2(g)$$

Calculate K_p for the equilibrium.

9. PCl_5 dissociates as follows:

$$PCl_5(g) = PCl_3(g) + Cl_2(g)$$

At 250° C and a pressure of 1 atm the density of the equilibrium mixture of the above gases was observed to be 2.695 g/liter. Calculate the degree of dissociation, α, of PCl_5, and K_p for the reaction.   *Ans.* $\alpha = 0.80$; $K_p = 1.78$

10. For the equilibrium

$$2 CO_2(g) = 2 CO(g) + O_2(g)$$
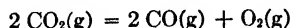
derive an expression for α, the degree of dissociation of CO_2, as a function of K_p for the reaction and the total pressure P. Assume that α is negligible compared to unity.

11. Derive for the equilibrium

$$C_2H_4(g) + H_2(g) = C_2H_6(g)$$
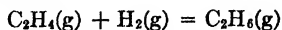
an expression for x, the number of moles of C_2H_6 present at equilibrium, in terms of a and b, the initial number of moles of C_2H_4 and H_2 respectively, the total pressure P, and K_p.

12. For the reaction

$$C_2H_5OH(l) + CH_3COOH(l) = CH_3COOC_2H_5(l) + H_2O(l)$$
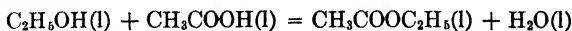
let a be the number of moles of alcohol present initially per mole of acid, and x the number of moles of acid esterified after equilibrium is established. Then from the following data

a	x
0.5	0.42
1.0	0.665
1.5	0.779

(a) calculate K_c in each case, and (b) from the average value of K_c find x when a is 0.1 mole.

13. At 3000° K and 1 atm, CO_2 is 40% dissociated into CO and O_2. (a) What will be the degree of dissociation if the pressure is raised to 2 atm? (b) What will be the degree of dissociation when a mixture of 50% CO_2 and 50% O_2 is heated to 3000° K, the pressure being 1 atm?   *Ans.* (a) 0.335; (b) 0.271

14. From the value of K_p in Table 1 estimate the percentage of NH_3 present at equilibrium at 450° C when the total pressure is 100 atm and the molar ratio of H_2 to N_2 is 3 to 1.

15. Repeat the calculation in the preceding problem, assuming that the molar ratio of H_2 to N_2 is 1 to 1 and the total pressure 10 atm.

16. At 30° C K_p in atmospheres for the dissociation

$$SO_2Cl_2(g) = SO_2(g) + Cl_2(g)$$

is 2.9×10^{-2}. Calculate the degree of dissociation when the total pressure is 1 atm.

17. At 25° C ΔF^0 for the reaction

$$N_2O_4(g) = 2\,NO_2(g)$$
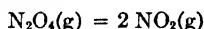
is $+ 1154$ cal. What is the degree of dissociation at 25° C when the total pressure is 10 atm?

18. In the preceding problem calculate the degree of dissociation when the total pressure is maintained at 10 atm, but a partial pressure of CO_2 equal to 5 atm is present at equilibrium.

19. For the dissociation of HI

$$2\,HI(g) = H_2(g) + I_2(g)$$

$K_c = 50.0$ at 448° C, and $K_c = 66.9$ at 350° C. Estimate (a) the heat of dissociation, and (b) the degree of dissociation, α, at 400° C. Assume that ΔH does not vary with temperature.   *Ans.* (a) $- 2660$ cal; (b) 0.995

20. For the reaction
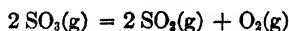
$$2\,SO_3(g) = 2\,SO_2(g) + O_2(g)$$
$$\Delta H_{298.2° K} = 43,800 \text{ cal} \qquad \Delta F^0_{298.2° K} = 31,800 \text{ cal}$$

Calculate (a) ΔF^0, and (b) the degree of dissociation, α, of SO_3 at 600° K and 0.5 atm pressure. Assume ΔH to be independent of temperature.

21. From the heat of formation and free energy of formation of NO and NO_2 given in the proper tables of the preceding chapters, find the degree of dissociation of NO_2 into NO and O_2 at 200° C and 1 atm pressure. Assume that ΔH is constant.

22. For the reaction

$$CO(g) + H_2O(g) = CO_2(g) + H_2(g)$$
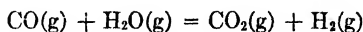
the following are the absolute entropies S^0, and heats of formation ΔH, at 25° C:

	S^0(eu)	ΔH (cal)
CO	47.32	− 26,600
H_2O	45.13	− 57,800
CO_2	51.08	− 94,500
H_2	31.23	—

Assuming that ΔH is constant with T, calculate K_p for the reaction at 600° K.

Ans. 29.0

23. By referring to Table 2 of the preceding chapter, predict which of the following reactions are thermodynamically possible in the standard state at 25° C:

   (a) $3 C_2H_2(g) = C_6H_6(g)$

   (b) $4 NH_3(g) + 5 O_2(g) = 4 NO(g) + 6 H_2O(g)$

   (c) $N_2O_4(g) = 2 NO(g) + O_2(g)$

   (d) $CO(g) + 2 H_2(g) = CH_3OH(l)$

24. For the reaction

$$CH_3\!\!-\!\!\underset{\underset{CH_3}{|}}{C}HOH(g) = CH_3\!\!-\!\!\underset{\underset{CH_3}{|}}{C}O(g) + H_2(g)$$
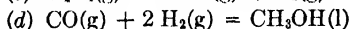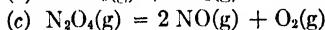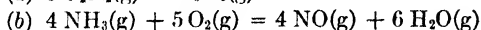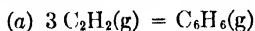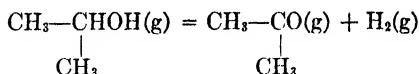
K_p at 457.4° K is 0.36, $\Delta C_p - 4.0$, and $\Delta H_{298.1° K} = 14,700$ cal. (a) Derive an expression for $\log_{10} K_p$ as a function of T, and (b) calculate K_p at 500° K.

Ans. (a) $\log_{10} K_p = -\dfrac{2950}{T} + 2.01 \log_{10} T + 0.656$; (b) $K_p = 1.42$

25. Using the Nernst approximation formula, estimate K_p at 350° K for the reaction:

$$I_2(s) + H_2S(g) = 2 HI(g) + S(s)$$
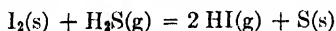
The heat for formation of H_2S is − 4800 cal/mole while that of HI is 5900 cal/mole.   *Ans.* 7.5×10^{-3}

26. Using the Nernst approximation formula, estimate the extent of dissociation of HI into H_2 and I_2 at 600° K.

Heterogeneous Equilibrium I: Mass Action and Distribution Laws

Introduction. In the preceding chapter attention was directed to equilibria occurring in homogeneous systems, and to the application of the law of mass action to such equilibria. The purpose of this and the next chapter is to develop the methods by which equilibria in heterogeneous systems may be handled.

Unlike a homogeneous system, in which only a single phase is present, a heterogeneous system is characterized by the presence of several distinct phases. Each of these phases is itself homogeneous, possesses definite physical and chemical properties, and is separated from every other phase by a surface boundary. As far as state of aggregation is concerned, each phase may be solid, liquid, or gas. Since all gases are miscible in all proportions, there can be only one gas phase in a system. On the other hand, in a system containing several liquids or solids, or solids and liquids, several phases may occur. Hence we may have heterogeneous systems composed of two or more solid phases, two or more liquid phases, or combinations of solid, liquid, and gas phases. As examples of heterogeneous systems may be mentioned two partially miscible liquids in contact, a liquid and its vapor, a solid and its melt, a solid and its vapor, a solution in presence of its vapor or solid solvent, a saturated solution in contact with saturating solid phase, a salt like barium carbonate in presence of its thermal decomposition products, barium oxide and carbon dioxide, or a system in which cupric oxide reacts with hydrogen to form copper and water vapor.

From the examples cited it is readily apparent that the number of equilibria possible in heterogeneous systems is very large. In general, the various equilibria encountered may be classified as being either physical in nature, such as the equilibrium between a liquid and its vapor, or chemical, such as dissociation of barium carbonate, although there are cases of mixed type, as we shall see. In approaching the problem of equilibria in heterogeneous systems three methods of attack are available: (1) the law of mass action, (2) the Nernst distribution law, and (3) the Gibbs phase rule. The particular method employed

depends on the nature of the problem at hand and the information sought. The law of mass action approach is applicable when quantitative information is desired on heterogeneous *chemical* equilibria involving gases; the Nernst distribution law is used in problems involving the solubility of a substance in two mutually insoluble solvents; while the phase rule is a general principle valuable for the definition of conditions obtaining in heterogeneous equilibria. These delineations of the sphere encompassed by each of these principles will become clearer as soon as specific instances of the application of each are given.

This chapter will be devoted to an exposition of heterogeneous equilibria from the standpoint of the law of mass action and the distribution law. The phase rule in its various aspects will be discussed in the next chapter.

Heterogeneous Equilibria and the Law of Mass Action. In applying the law of mass action to heterogeneous chemical equilibria, *the presence of pure solid or liquid phases is completely disregarded in writing the expression for the equilibrium constant.* To understand why this is so, consider as an example the heterogeneous equilibrium

$$CuO(s) + H_2(g) = Cu(s) + H_2O(g) \tag{1}$$
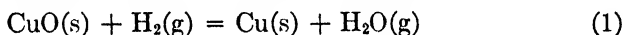
According to equation (16) of Chapter XII, the thermodynamic equilibrium constant K_a for this reaction should be

$$K_a = \frac{a_{Cu} a_{H_2O}}{a_{CuO} a_{H_2}} \tag{2}$$

or, since for a gas the activity is equal to the fugacity f,

$$K_a = \frac{a_{Cu} f_{H_2O}}{a_{CuO} f_{H_2}} \tag{3}$$

However, it was pointed out in Chapter XI that *the activity of a pure solid or liquid may be taken as unity at all temperatures up to fairly high pressures.* Then $a_{Cu} = a_{CuO} = 1$, and

$$K_a = \frac{f_{H_2O}}{f_{H_2}} \tag{4}$$

When the gases involved may be considered to behave ideally, the fugacity reduces to the pressure, and equation (4) becomes

$$K_p = \frac{P_{H_2O}}{P_{H_2}} \tag{5}$$

Thermodynamic reasoning indicates, therefore, that the equilibrium constant for a heterogeneous reaction should contain only the fugacities or pressures of the gaseous constituents, and should not include in any way terms for either pure solids or pure liquids.

Equation (5) may also be arrived at from a different point of view. Before considering this alternate method of deduction, it is necessary to show that for a pure solid or liquid the vapor pressure may be used as a measure of the "active mass" of the substance in the solid or liquid state. As the student is aware, many solid substances, such as iodine or camphor for instance, exhibit appreciable and measurable sublimation pressures. This pressure is due to the evaporation of a portion of the solid, is independent of the amount of solid present, and at any given temperature is constant as long as any solid phase remains in contact with the vapor. On the other hand, there are many solids over which no such vapor tension can be detected at ordinary temperatures. Still, the fact that sublimation pressures can be made manifest over these substances at higher temperatures is an indication that even these solids must be considered to possess sublimation pressures, though such pressures may be so small as to be beyond experimental observation with the means at our disposal. These same arguments and conclusions are equally applicable to liquids. Consequently, whether measurable or not, it is convenient for theoretical purposes to ascribe to every solid or liquid a vapor pressure which for each substance is constant and characteristic at a given temperature.

For equilibrium to exist between a solid or a liquid and its vapor both the solid and its vapor or the liquid and its vapor must have equal "active masses." Since the pressure of the vapor is a measure of its "active mass," and since this pressure is constant for any given solid or liquid at any given temperature, it must follow that the "active mass" of a solid or liquid should also be constant at constant temperature. This deduction is amply supported by experimental evidence in the chemical literature which shows that at constant temperature the "active mass" of a solid or liquid is fixed, i.e., is completely independent of the amount of solid or liquid present. We are led to the conclusion, therefore, that the sublimation pressures of solids or the vapor pressures of liquids may be used as a measure of the "active masses" of pure solids and liquids involved in chemical reactions.

If this be the case, the equilibrium constant for the reactions given in equation (1) may be written as

$$K_p = \frac{P_{\text{Cu}}P_{\text{H}_2\text{O}}}{P_{\text{CuO}}P_{\text{H}_2}} \qquad (6)$$

where P_{Cu} and P_{CuO} are the sublimation pressures of the two solids. Since these pressures are constant at any given temperature, they may be incorporated into K_p to yield a new constant K_p', thus

$$K_p' = K_p \frac{P_{\text{CuO}}}{P_{\text{Cu}}} = \frac{P_{\text{H}_2\text{O}}}{P_{\text{H}_2}} \qquad (7)$$

K'_p in equation (7) is identical with K_p in equation (5), although the two were arrived at from different considerations. As may be surmised from the derivation of equation (5), there is a tacit assumption buried in equation (7), namely, that the gases behave ideally. Both K_p and K'_p for heterogeneous reactions are generally referred to as *condensed* equilibrium constants.

Examples of Heterogeneous Equilibria. Several typical examples of heterogeneous equilibria are given below to illustrate both some of the methods of studying such equilibria, and the application of the law of mass action to these.

(a) *The Dissociation of Cupric Oxide.* At elevated temperatures cupric oxide dissociates into cuprous oxide and oxygen according to the reaction

$$4\,CuO(s) = 2\,Cu_2O(s) + O_2(g) \tag{8}$$
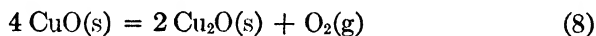
Since oxygen is the only gaseous constituent involved in this equilibrium, the condensed equilibrium constant K'_p for this reaction, in view of the above discussion, should be

$$K'_p = P_{O_2} \tag{9}$$

i.e., the pressure of oxygen above a mixture of cupric oxide and cuprous oxide should be constant at each given temperature. This conclusion is verified by the experimental results of F. Hastings Smith and H. R. Robert.[1] These investigators placed a charge of pure cupric oxide in a silica tube, evacuated the system, and then heated the tube to the desired temperature in an electric furnace. The equilibrium pressure of oxygen developed as a result of dissociation of the cupric oxide was read on a manometer attached to the silica tube. To make sure that the pressures read were those at true equilibrium, the latter was approached from both lower and higher temperatures, and at each temperature some gas was removed and the equilibrium pressure permitted to reestablish itself. In this manner Hastings Smith and Robert proved that at each temperature the pressure of oxygen is constant in accord with equation (9). Some of their results at various temperatures are given in Table 1.

TABLE 1

DISSOCIATION PRESSURES OF CuO AT VARIOUS TEMPERATURES

$t°\,C$	$K'_p = P_{O_2}$ (mm Hg)
900	12.5
940	29.2
980	65.0
1020	137.7
1060	278.0
1080	388.0

[1] F. Hastings Smith and H. R. Robert, J. Am. Chem. Soc., **42**, 2582 (1920)

(b) *The Carbon Disulfide Equilibrium.* When gaseous sulfur is passed over carbon at high temperatures, carbon disulfide is formed according to the equation

$$C(s) + S_2(g) = CS_2(g) \tag{10}$$
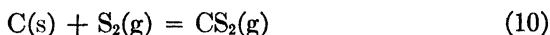
For this reaction the condensed equilibrium constant is given by

$$K'_p = \frac{P_{CS_2}}{P_{S_2}} \tag{11}$$

To study this equilibrium F. Koref [1] employed a dynamic method. Nitrogen gas saturated with sulfur vapor was passed over finely divided carbon kept at the desired temperature in an electric furnace. The exit gases were cooled rapidly to "freeze" the equilibrium, i.e., to prevent the shift of equilibrium, the sulfur and carbon disulfide condensed out, while the nitrogen was collected in a gasometer. The equilibrium quantities of carbon disulfide and sulfur were then determined by weighing.

From the data thus obtained the equilibrium constant was calculated as follows. If it is assumed that the sulfur and carbon disulfide vapors behave ideally, the volume in cubic centimeters of each of these at equilibrium at temperature T and total pressure P is

$$V_{S_2} = \frac{n_{S_2}RT}{P} \qquad V_{CS_2} = \frac{n_{CS_2}RT}{P}$$

where n_{S_2} and n_{CS_2} are the numbers of moles of these substances present in the condensates from the equilibrium mixture. If we let V be the total volume of gases at equilibrium, i.e., the sum of the volumes of sulfur, carbon disulfide, and nitrogen, then the mol fractions of the first two substances are $N_{S_2} = V_{S_2}/V$ and $N_{CS_2} = V_{CS_2}/V$, and hence, according to Dalton's law,

$$P_{S_2} = N_{S_2}P = \left(\frac{V_{S_2}}{V}\right)P \qquad P_{CS_2} = N_{CS_2}P = \left(\frac{V_{CS_2}}{V}\right)P$$

Substituting these expressions for the partial pressures into equation (11), we find

$$\begin{aligned} K'_p &= \frac{P_{CS_2}}{P_{S_2}} \\ &= \frac{\dfrac{V_{CS_2}P}{V}}{\dfrac{V_{S_2}P}{V}} \\ &= \frac{V_{CS_2}}{V_{S_2}} \end{aligned} \tag{12}$$

[1] F. Koref, Z. anorg. Chem., **66**, 73 (1910).

namely, the equilibrium constant should equal the ratio of the volumes of the two substances at equilibrium. Table 2 lists some of Koref's data at 1009° C and the values of K'_p calculated from these by means of equation (12). As may be seen from the third column, the constancy exhibited by K'_p is satisfactory.

<div align="center">

TABLE 2

K'_p AT 1009° C FOR THE REACTION C(s) + S$_2$(g) = CS$_2$(g)

</div>

V_{CS_2}(cc)	V_{S_2}(cc)	K'_p
458	84	5.45
607	109	5.57
738	130	5.68
814	142	5.73
1164	207	5.62
2057	371	5.54

(c) **The Dissociation of Ammonium Carbamate.** Ammonium carbamate (NH$_2$COONH$_4$) dissociates even at room temperature as follows

$$NH_2COONH_4(s) = 2\,NH_3(g) + CO_2(g) \tag{13}$$
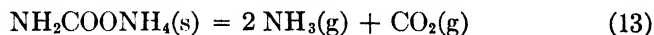
The condensed equilibrium constant for this reaction is, therefore,

$$K'_p = P^2_{NH_3} P_{CO_2} \tag{14}$$

In studying this equilibrium by a static method T. R. Briggs and V. Migidichian [1] introduced solid ammonium carbamate along with definite quantities of ammonia and carbon dioxide into an evacuated glass vessel to which was attached a manometer. The apparatus was immersed then in a water thermostat, the mixture was allowed to reach equilibrium, and the total pressure at equilibrium was read on the manometer.

If we consider the experiments where only ammonia was introduced initially along with the ammonium carbamate, and if we let e_1 be the initial pressure of ammonia and P the total pressure at equilibrium, then the increase in pressure due to dissociation of the solid is $(P - e_1)$. From the stoichiometry of the reaction it follows that of this increase two-thirds must be due to the formation of ammonia, one-third to formation of carbon dioxide. Consequently,

$$P_{NH_3} = \frac{2}{3}(P - e_1) + e_1 = \left(\frac{2P + e_1}{3}\right)$$

$$P_{CO_2} = \frac{1}{3}(P - e_1)$$

[1] Briggs and Migidichian J. Phys. Chem., **28**, 1121 (1924).

and equation (14) becomes

$$K'_p = \left(\frac{2\,P + e_1}{3}\right)^2\left(\frac{P - e_1}{3}\right)$$
$$= \frac{(2\,P + e_1)^2(P - e_1)}{27} \tag{15}$$

Table 3, taken from the paper mentioned, presents typical data obtained during a series of experiments at 30° C. In this table e_1 and P have the significance ascribed to them above, P_{NH_3} and P_{CO_2} are the equilibrium pressures of the two gases, and K'_p are the equilibrium constants calculated according to (15). All pressures are given in mm Hg. From equation (15) it is apparent that when $e_1 = 0$, i.e., when no ammonia is present initially, the expression for K'_p reduces to

$$K'_p = \frac{(2\,P)^2 P}{27}$$
$$= \frac{4\,P^3}{27} \tag{16}$$

Other instances of heterogeneous equilibria which can be handled in the manner described are the thermal dissociations of metal carbonates, such as those of calcium, barium, strontium, lead, zinc, copper, manganese, and iron, dissociations of oxides, such as silver oxide, mercuric oxide, and cobalt oxide, dissociations of halides such as copper bromide,

TABLE 3

DISSOCIATION OF AMMONIUM CARBAMATE AT 30° C

e_1	P	P_{NH_3}	P_{CO_2}	K'_p
0	125.0	83.3	41.7	286×10^3
13.6	124.9	87.8	41.7	281
27.3	125.4	92.7	37.1	277
52.5	129.5	103.8	32.7	289
141.1	174.2	163.2	11.0	293
168.6	194.2	185.7	8.5	293

silver bromide, silver iodide, and dissociations of such sulfides as ammonium hydrosulfide, ammonium sulfide, and bismuth sulfide. Still further examples are the oxidation of carbon to carbon dioxide, reduction of carbon dioxide by carbon to carbon monoxide, the "water gas" reaction,

$$H_2O(g) + C(s) = H_2(g) + CO(g)$$

and the reduction of metal oxides by carbon monoxide, as

$$Fe_2O_3(s) + 3\,CO(g) = 3\,CO_2(g) + 2\,Fe(l)$$
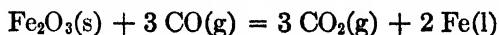
the reaction occurring in the blast furnace during the preparation of iron from hematite ore.

Effect of Pressure on Heterogeneous Equilibria. As in homogeneous equilibria, the influence of pressure on heterogeneous equilibria can be predicted by means of the Le Chatelier principle provided that two conditions are fulfilled, namely, (1) that the system is actually at equilibrium, and (2) that the pressure on the system can be changed. To appreciate the first of these conditions it is essential to realize that the equilibrium constant expression for a given heterogeneous reaction is valid only *as long as all the condensed phases participating in the equilibrium are present at equilibrium.* In order to prove this point, consider again the equilibrium

$$4\,CuO(s) = 2\,Cu_2O(s) + O_2(g) \qquad (17)$$
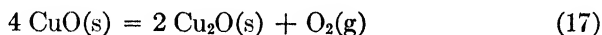
According to the law of mass action, the pressure of oxygen at equilibrium in this system should be constant at any given temperature. When the equilibrium is approached from the cupric oxide side, and when sufficient cupric oxide is taken to supply the given pressure of the gas in the volume in which it is placed, all the phases participating in the equilibrium are present, and the demands of the law of mass action can be satisfied. Suppose, however, that the equilibrium is approached from the other side, namely, with oxygen and cuprous oxide, and suppose that the initial pressure of oxygen is considerably above the equilibrium pressure. Then cuprous oxide will interact with the oxygen to form cupric oxide, and this reaction will proceed until the pressure of the gas has been reduced to the equilibrium pressure. But, this reaction can go this far only if sufficient cuprous oxide is present to react with all the excess oxygen. If such is not the case, all the cuprous oxide will be converted to cupric oxide, and still an excess of oxygen will remain to yield a pressure higher than that demanded by the equilibrium constant. Under such conditions no equilibrium is possible in the system, since no cuprous oxide is present, and hence the oxygen pressure may assume any value dependent on the amount of it present.

These conclusions apply to all heterogeneous equilibria. It must be remembered, therefore, that sufficient amounts of the solid phases involved in the equilibria must be present to permit regulation and adjustment of the conditions in the system necessary for the establishment of true equilibrium. Otherwise true equilibrium is impossible, the equilibrium constant principle does not hold, and the Le Chatelier principle is not applicable.

The second condition under which the Le Chatelier principle cannot be used is when the pressure of the system at equilibrium cannot be changed at constant temperature. An example of this situation is again

the cupric oxide-cuprous oxide-oxygen reaction. An attempt to increase the oxygen pressure at equilibrium by adding oxygen to the system or by compression will result in a reaction of cuprous oxide with the gas until the constant pressure of the latter demanded by the equilibrium constant is reestablished. Similarly, an attempt to reduce the pressure by removal of oxygen or by increasing the volume of the system will result in a dissociation of cupric oxide until the constant equilibrium pressure is again attained. Therefore, in a system in which a single gaseous constituent is in equilibrium with several solid phases it is impossible to change the pressure as long as all the solid phases exist at equilibrium. There can be no question here of applying the Le Chatelier principle to predict the effect of pressure change.

In all other instances the Le Chatelier principle may be employed to foretell the direction in which the point of equilibrium will be shifted by a change in pressure. Whenever a reaction proceeds with a decrease in volume, increase of pressure will favor the products, while for reactions occurring with an increase in volume, higher pressures will favor the reactants. In the special case of no volume change, the extent of reaction will be uninfluenced by the total equilibrium pressure. In considering the change in volume accompanying a reaction, the volumes of all condensed phases may be disregarded, since they are negligibly small compared to those of the gases involved.

An interesting equilibrium in this connection is the dissociation of ammonium carbamate, equation (13). When the equilibrium is established in absence of any excess products, K_p' is given by equation (16), P is constant at any given temperature, and the equilibrium behaves like that of cupric oxide-cuprous oxide. When one of the products is present initially, however, the total equilibrium pressure is no longer constant but depends on the amount of product started with, as may be seen from either equation (15) or Table 3 for the case of excess ammonia. In all instances initial introduction of ammonia into the system leads to a higher equilibrium pressure, and, therefore, the dissociation of ammonium carbamate should be more inhibited as the pressure increases. That such is actually the case is apparent from a comparison of the second and fourth columns of the table. The same conclusion could have been arrived at by utilizing the law of mass action to predict the effect of presence of initial gaseous products on the extent of reaction.

Effect of Temperature on Heterogeneous Equilibria. The variation of equilibrium constants of heterogeneous reactions with temperature is given by the same equations as that of homogeneous reactions, namely, the differential equations (41) and (46) of Chapter XII and the integrated forms which correspond to these. The validity of these equa-

tions may again be tested by plotting log K'_p vs. $1/T$ and observing whether a straight line is obtained. Figure 1 shows such a plot of the data presented in Table 1 for the thermal dissociation of cupric oxide.
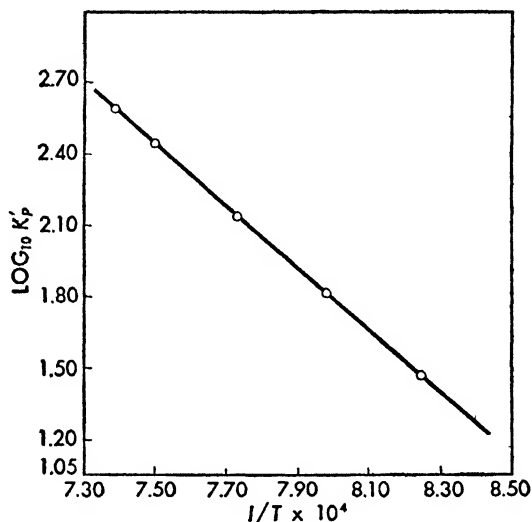
Fig. 1. Plot of $\text{Log}_{10} K'_p$ vs. $1/T$ for the Reaction: $4\,CuO(s) = 2\,Cu_2O(s) + O_2(g)$

The straight line thus obtained has a slope of $-13,200$, and since this slope is equal to $-\Delta H/2.303\,R$, it follows that the heat of the dissociation reaction is:

$$-\frac{\Delta H}{2.303\,R} = -13,200$$
$$\Delta H = 4.57 \times 13,200$$
$$= 60,300 \text{ cal}$$

Equilibria in Hydrates. A particularly interesting and important type of heterogeneous equilibrium is that exhibited by hydrates, or in general solvates, of various substances. Under proper conditions these hydrates are found to dissociate into lower hydrates, or the anhydrous substance, and water vapor which establishes a definite vapor pressure above the solid phases. Thus $Na_2HPO_4 \cdot 12\,H_2O$ dissociates into $Na_2HPO_4 \cdot 7\,H_2O$, the latter into $Na_2HPO_4 \cdot 2\,H_2O$, the dihydrate into Na_2HPO_4. The equilibria attending these dissociations are:

$$Na_2HPO_4 \cdot 12\,H_2O(s) = Na_2HPO_4 \cdot 7\,H_2O(s) + 5\,H_2O(g) \quad (18)$$
$$Na_2HPO_4 \cdot 7\,H_2O(s) = Na_2HPO_4 \cdot 2\,H_2O(s) + 5\,H_2O(g) \quad (19)$$
$$Na_2HPO_4 \cdot 2\,H_2O(s) = Na_2HPO_4(s) + 2\,H_2O(g) \quad (20)$$
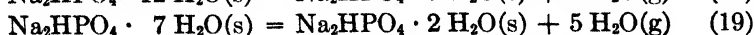
These heterogeneous equilibria can be treated by the methods described. It is readily apparent that the equilibrium constants of all such dissociations are given simply by

$$K'_p = P^x \tag{21}$$

where x is the number of moles of vapor resulting from the dissociation of the hydrate, i.e., 5 for equations (18) and (19) and 2 for equation (20), while P is the vapor pressure or aqueous tension above the pair of solid phases. From equation (21) it follows that, since P^x is constant for a particular dissociation, P must also be constant, and therefore at any given temperature the vapor pressure above any *hydrate pair* must be constant as long as both phases are present. This conclusion is in accord with observation. In the presence of any given hydrate and its lower dissociation product the pressure is found to be definite and characteristic of the particular hydrate pair. This is not true, however, when only a single solid phase is present. Under such conditions the vapor pressure can vary within certain ranges, as we shall see, and the deductions made above do not apply. It is erroneous, therefore, to speak of the vapor pressure of a hydrate: the vapor pressure is constant only for a hydrate pair or a hydrate and its anhydride.

Fig. 2. Vapor Pressures of the Hydrates of Na₂HPO₄ at 25° C

In Table 4 are given the vapor pressures in mm Hg at 25° C for the hydrate pairs of several salts. It will be observed that in general the vapor pressure is highest for the pair richest in water and decreases as the water content of the solid phase decreases. A clearer appreciation of the relation between the pressures of the various hydrated forms of a substance at a particular temperature can be obtained by plotting the pressure against the number of moles of water in the solid phase, n. Such a plot for disodium phosphate is shown in Fig. 2. The horizontal portions indicate the values of n over which the particular vapor pressures will remain constant, i.e., between $n = 0$ and $n = 2$, between $n = 2$ and $n = 7$, and $n = 7$ and $n = 12$. The vertical portions, on the other hand, give the *vapor pressure ranges* over which the pure solid phases (anhydrous salt, 2 H₂O, 7 H₂O, and 12 H₂O) are stable. Thus it is seen that between zero and 9.80 mm Hg water vapor pressure anhy-

TABLE 4

AQUEOUS VAPOR PRESSURE OF HYDRATE PAIRS AT 25° C

(In mm Hg)

Hydrate Pair	P	Reference
$MgSO_4 \cdot 7 H_2O - MgSO_4 \cdot 6 H_2O$	11.5	
$6 H_2O - 5 H_2O$	9.8	Foote and Scholes, J. Am.
$5 H_2O - 4 H_2O$	8.8	Chem. Soc., **33**, 1309 (1911)
$4 H_2O - H_2O$	4.1	
$H_2O - MgSO_4$	1.0	
$CuSO_4 \cdot 5 H_2O - CuSO_4 \cdot 3 H_2O$	7.80	
$3 H_2O - H_2O$	4.7	Wilson, *ibid.*, **4**, 704 (1921)
$H_2O - CuSO_4$	0.8(?)	
$Na_2HPO_4 \cdot 12 H_2O - Na_2HPO_4 \cdot 7 H_2O$	19.13	Noyes and Westbrook, *ibid.*,
$7 H_2O - 2 H_2O$	14.51	**43**, 726 (1921)
$2 H_2O - Na_2HPO_4$	9.80	

drous disodium phosphate does not combine with water vapor. As soon as the latter pressure is reached, however, some $Na_2HPO_4 \cdot 2 H_2O$ is formed, and the pressure remains constant as long as any unconverted anhydrous salt is present. As soon as all of the disodium phosphate has been converted to the dihydrate it is found that the aqueous tension can be varied between 9.80 and 14.51 mm Hg without formation of the heptahydrate; namely, the range of stability of the dihydrate is from 9.80 to 14.51 mm Hg vapor pressure. At 14.51 mm Hg the heptahydrate phase begins to form, and the pressure again becomes constant, this time at 14.51 mm, until all of the dihydrate has been converted to the heptahydrate. The latter in turn remains stable between this pressure and 19.13 mm, at which pressure the 12 H_2O begins to form. Once the solid phase has been converted to the 12 H_2O, the pressure may be increased up to the pressure of the saturated solution of disodium phosphate in water. On reaching this pressure some of the solid will be found to dissolve to a solution in equilibrium with solid $Na_2HPO_4 \cdot 12 H_2O$.

On dehydration the phenomena described will be found to take place in reverse order, i.e., the 12 H_2O will break down to the 7 H_2O, the 7 H_2O to 2 H_2O, etc. From this discussion it is apparent that any particular hydrate can be preserved only as long as the aqueous tension lies between the limits of stability exhibited by the hydrate at the given temperature. Outside these limits the hydrate will either dissociate to a lower hydrate or the anhydrous salt, or will take on water to form a water richer phase, depending on whether the vapor pressure is below or above the prescribed limits.

This description of the behavior of the hydrates of disodium phosphate at various aqueous tensions is typical of that exhibited by all hydrates at a particular temperature, and may be taken as a general definition of

the behavior of such substances. The same considerations apply also to other solvates, such as alcoholates and ammoniates of various salts.

The vapor pressures of hydrate and other solvate pairs increase with rise in temperature, as may be seen from Table 5, in a manner readily deducible from equation (21). Since $K'_p = P^x$, then

$$\frac{d \ln P^x}{dT} = \frac{\Delta H}{RT^2} \tag{22}$$

and therefore,

$$\frac{d \ln P}{dT} = \frac{\Delta H}{xRT^2} \tag{23}$$

On integration, assuming ΔH constant, this leads to

$$\log_{10} P = \left(\frac{-\Delta H}{2.303\,xR}\right)\frac{1}{T} + C \tag{24}$$

Inspection of equation (24) shows it to be identical with the Clausius-Clapeyron equation except that the slope of $\log_{10} P$ vs. $1/T$, instead of being $(-\Delta H/2.303\,R)$, is given by $(-\Delta H/2.303\,xR)$. From the slope of such a plot the heat of dissociation of a hydrate or solvate readily follows as

$$\Delta H = -2.303\,xR \text{ (slope)} \tag{25}$$

For the special instance of $x = 1$ equation (24) becomes identical with the Clausius-Clapeyron equation.

TABLE 5

VAPOR PRESSURES OF SOME HYDRATE PAIRS AT VARIOUS TEMPERATURES
(P in mm Hg)

Hydrate Pair	0° C	15°	20°	25°	30°
$Na_2HPO_4 \cdot 12\,H_2O - 7\,H_2O$ *	2.66	8.95	12.93	19.18	27.05
$SrCl_2 \cdot 6\,H_2O - 2\,H_2O$ †	1.23	3.99	8.37	—	—
$Na_2SO_4 \cdot 10\,H_2O - Na_2SO_4$ †	2.77	9.21	19.20	—	—

* Partington and Winterton, J. Chem. Soc., **132**, 635 (1930); Baxter and Cooper, J. Am. Chem. Soc., **46**, 927 (1924).
† Baxter and Lansing, J. Am. Chem. Soc., **42**, 419 (1920).

Heterogeneous Equilibria Involving Ions. All the heterogeneous equilibria to which the law of mass action has been applied so far involved gases and condensed phases. This law can be applied also to heterogeneous equilibria involving condensed phases and ions in solution, but a discussion of such equilibria will be postponed to the chapter on Ionic Equilibria.

The Distribution Law. Mercuric bromide is soluble in both water and benzene. When a solution of mercuric bromide in water is shaken with benzene, which is immiscible with water, it is found that the mercuric bromide distributes itself between the water and benzene layers in such a way that at equilibrium the *ratio of the concentrations* of mercuric bromide in the two layers *is a constant at any given temperature.* This distribution of a solute between two immiscible or only slightly miscible solvents can be accomplished with any solute for which a pair of immiscible solvents can be found which will dissolve the substance.

The tendency of a substance to distribute itself between immiscible solvents so that the concentration ratio is a constant is a direct consequence of the thermodynamic requirements for equilibrium. To show this, consider a pair of immiscible solvents in contact, A and B, both containing the same substance in solution. According to equation (28) of Chapter XI, the free energy of the solute in liquid A can be represented by

$$F_A = F_A^0 + RT \ln a_A \qquad (26)$$

where F_A is the free energy of the solute in A, F_A^0 the standard free energy of the solute in the same liquid, i.e., the value of F_A for $a = 1$, while a_A is the activity of the solute in solvent A. Similarly the free energy of the solute in the second liquid can be written as

$$F_B = F_B^0 + RT \ln a_B \qquad (27)$$

where all the quantities have the same significance as in equation (26), except that they refer now to liquid B. Since for equilibrium between the layers $\Delta F = F_B - F_A$ must be zero, or that $F_B = F_A$, it follows from equations (26) and (27) that

$$F_B = F_A$$
$$F_B^0 + RT \ln a_B = F_A^0 + RT \ln a_A$$
$$RT \ln \frac{a_B}{a_A} = F_A^0 - F_B^0$$
$$\ln \frac{a_B}{a_A} = \frac{F_A^0 - F_B^0}{RT} \qquad (28)$$

However, at any given temperature F_B^0 and F_A^0 are constants for a given substance in the particular solvents. Hence,

$$\ln \frac{a_B}{a_A} = k$$

and therefore,

$$\frac{a_B}{a_A} = K \qquad (29)$$

Equation (29) is a mathematical statement of the *distribution law*, which states that a substance will distribute itself between two solvents until at equilibrium the *ratio* of the activities of the substance in the two layers is constant at any given temperature, irrespective of the absolute values of either a_B or a_A. When the solutions are dilute, or when the solute behaves ideally, the activity is essentially equal to the concentration C, and equation (29) reduces to the more common form

$$\frac{C_B}{C_A} = K \tag{30}$$

The constant K is generally referred to as either the *distribution* or *partition coefficient* of the solute between the two solvents.

The applicability of the simplified distribution law, equation (30), may be judged from the data given in Table 6 for the distribution at 25° C of iodine between carbon tetrachloride and water, mercuric bromide between water and benzene, and boric acid between water and amyl alcohol. The essential constancy of the distribution coefficients for the various low concentrations given shows that in dilute solutions equation (30) is valid. However, the last two values for iodine and the last for boric acid indicate that in more concentrated solutions activities must be employed instead of concentrations to obtain a true constant for the partition coefficient. Furthermore, K depends on the nature of the solute and the liquids involved. Other factors affecting the magnitude of this constant are the temperature and the manner in which the

TABLE 6

DISTRIBUTION COEFFICIENTS AT 25° C

(C in moles/liter)

I_2 between H_2O and CCl_4 *			$HgBr_2$ between H_2O and C_6H_6 †			H_3BO_3 between H_2O and Amyl Alcohol ‡		
C_{H_2O}	C_{CCl_4}	$K = \dfrac{C_{H_2O}}{C_{CCl_4}}$	C_{H_2O}	$C_{C_6H_6}$	$K = \dfrac{C_{H_2O}}{C_{C_6H_6}}$	C_{H_2O}	C_A	$K = \dfrac{C_{H_2O}}{C_A}$
0.000322	0.02745	0.0117	0.00320	0.00353	0.90	0.02602	0.00805	3.24
0.000503	0.0429	0.0117	0.00394	0.00436	0.90	0.05104	0.01545	3.31
0.000763	0.0654	0.0117	0.00634	0.00715	0.89	0.1808	0.0540	3.35
0.00115	0.101	0.0114	0.00953	0.01074	0.89	0.3012	0.0857	3.52
0.00134	0.1196	0.0112	0.01147	0.01303	0.88			
			0.017	0.0194	0.88			

* Jakowkin, Z. physik. Chem., **18**, 585 (1895).
† Sherrill, *ibid.*, **43**, 730 (1903).
‡ Mueller and Abegg, *ibid.*, **57**, 514 (1907).

constant is written, i.e., C_A/C_B or C_B/C_A. As there is no generally agreed method of writing the partition coefficient, it is essential, in order to avoid ambiguity, to specify in each instance the equation for the constant and the temperature to which it applies.

Other examples of distribution which may be cited are the distributions of iodine between water and carbon disulfide, chloroform, and ethylene glycol, bromine between water and carbon disulfide or bromoform, hydrogen peroxide between water and various organic solvents, and phenol between water and amyl alcohol. Data and references on these distributions and many others are tabulated in the International Critical Tables, Vol. III, pp. 419–434.

Changes in Solute and the Distribution Law. Walter Nernst [1] first called attention to the fact that the above considerations concerning the distribution law and the relations deduced may be expected to be valid only if the solute on dissolving undergoes no change such as dissociation or association. If a solute does dissociate into ions or simpler molecules or if it associates into more complex molecules, the distribution law cannot be expected to apply to the total concentrations in the two phases, but only to the concentrations of the molecular species *common* to both phases. Thus, if a substance A dissolves in one solvent without any change in molecular form, and in another with partial association into, say, A_2, the partition coefficient for the distribution will not be given by the ratio of the total concentrations in the two phases, but rather by the total concentration in the first solvent divided by the concentration of unassociated molecules in the second solvent; i.e., by the ratio of the concentrations of the molecules having identical molecular weights in the two solvents. This limitation to the applicability of the distribution law was pointed out in conjunction with the use of Henry's law for gases that react with the solvent to form new molecular or ionic species.

In order to illustrate how the distribution law may be utilized in handling more complicated cases, consider the distribution of benzoic acid between water and chloroform, for which data are given in Table 7. Column 1 gives the total concentration of benzoic acid in the water layer, column 2 the same data for the chloroform layer at equilibrium, while column 3 shows the ratio of the two concentrations. Obviously there is no semblance of constancy in the ratio of total concentrations of benzoic acid in the two solvents. However, the observed results can be explained on the basis that the acid is partially *dissociated* in water into benzoate and hydrogen ions, and *associated* in chloroform into double molecules $(C_6H_5COOH)_2$. Since the distribution law can be applied here only to the common molecular species in the two phases, i.e., the single

[1] Nernst, Z. physik. Chem., **8**, 110 (1891).

molecules of the acid, the equilibria occurring in both phases must be considered. To do this, let

C_W = total concentration of benzoic acid in water (moles per liter)
C_C = total concentration of benzoic acid in $CHCl_3$ (expressed in moles per liter of single molecules)
α = degree of dissociation of the acid in water
m = concentration of single molecules of the acid in $CHCl_3$.

In terms of these quantities the equilibrium constants for the several equilibria in the system are given by:

Water:
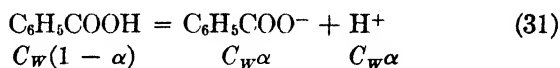
$$C_6H_5COOH = C_6H_5COO^- + H^+ \qquad (31)$$
$$C_W(1 - \alpha) \qquad C_W\alpha \qquad C_W\alpha$$
$$K_i = \frac{(C_W\alpha)^2}{C_W(1 - \alpha)}$$
$$= \frac{C_W\alpha^2}{(1 - \alpha)} \qquad (32)$$

Chloroform:
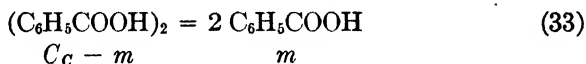
$$(C_6H_5COOH)_2 = 2\ C_6H_5COOH \qquad (33)$$
$$C_C - m \qquad m$$
$$K_1 = \frac{m^2}{C_C - m} \qquad (34)$$

Distribution:
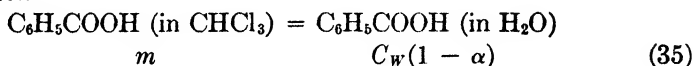
$$C_6H_5COOH\ (\text{in } CHCl_3) = C_6H_5COOH\ (\text{in } H_2O)$$
$$m \qquad C_W(1 - \alpha) \qquad (35)$$
$$K = \frac{C_W(1 - \alpha)}{m} \qquad (36)$$

Solving for m in equation (36) and substituting into equation (34), we obtain

$$K_1 = \frac{m^2}{C_C - m} = \frac{\left[\dfrac{C_W(1 - \alpha)}{K}\right]^2}{\left[C_C - \dfrac{C_W(1 - \alpha)}{K}\right]} \qquad (37)$$

Equation (37) involves three unknowns, K_1, K, and α. However, α for each value of C_W is available through equation (32) from the known value of K_i for benzoic acid in water, 6×10^{-5}. Furthermore, since K_1 should be a constant independent of concentration, it is possible to take two sets of values of C_W and C_C and the corresponding values of α and solve for K in equation (37). Once K is known, it may be employed

to find m through equation (36); and, as a check on the complete argument, m along with C_C may be used through equation (34) to see whether K_1, the dissociation constant for double molecules in benzene, comes out a true constant. In Table 7, columns 4 to 8, are given the concentrations of the various molecular and ionic species as deduced by W. S. Hendrixon[1] in the manner described. As may be seen from the last column of the table, K_1 does come out to be a good constant, and hence it is established that the explanation advanced does account for the observed distribution behavior. The actual distribution constant for single molecules K is 0.442 at 40° C.

Other systems to which such analysis has been applied with success are the distribution of benzoic acid between water and benzene, salicylic acid between water and benzene or chloroform, and acetic acid between water and various organic solvents.

TABLE 7

DISTRIBUTION OF BENZOIC ACID BETWEEN WATER AND CHLOROFORM AT 40° C
(C in moles/liter. $K = 0.442$)

C_W	C_C	$\dfrac{C_W}{C_C}$	α	$C_W(1-\alpha)$	m	$C_C - m$	$K_1 = \dfrac{m^2}{C_C - m}$
0.00211	0.00721	0.292	0.1548	0.00178	0.00404	0.00317	0.0052
0.00268	0.01084	0.247	0.1385	0.00231	0.00523	0.00561	0.0049
0.00353	0.01686	0.210	0.1220	0.00310	0.00701	0.00985	0.0050
0.00725	0.05700	0.127	0.0869	0.00662	0.01497	0.04203	0.0053
0.01272	0.16733	0.076	0.0660	0.01188	0.02687	0.14046	0.0052

Distribution Coefficients and Temperature. Distribution coefficients, like other equilibrium constants, may be expected to vary with temperature, as is actually the case. Thus K for the distribution of benzoic acid between water and chloroform is 0.564 at 10° C and 0.442 at 40° C. The variation of these constants with temperature may be represented by equation (47), Chapter XII, where ΔH is now the heat of transfer per mole of the solute from one solution to the other.

Applications of the Distribution Law. The distribution law can be and has been applied to the study of a number of problems of both theoretical and practical interest, such as extraction, analysis, and determination of equilibrium constants. Extraction is a subject of great importance both in the laboratory and in industry. In the laboratory occasion frequently arises for the removal of a dissolved substance from, say, a water solution with solvents such as ether, chloroform, carbon

[1] W. S. Hendrixon, Z. anorg. Chem., **13**, 73 (1897). Hendrixon gave his concentrations in grams per 200 cc of solution. In Table 7 these have all been converted to moles of single molecules per liter.

tetrachloride, or benzene. Again, in industry extraction is used to remove various undesirable constituents of a product, such as harmful ingredients in petroleum oils, by treating the product with an immiscible solvent in which the impurity is also soluble. In all such processes it is important to know how much solvent and how many treatments are necessary in order to accomplish a particular degree of separation.

When a substance distributes itself between two solvents without the complications of association, dissociation, or reaction with the solvent, it is possible to calculate the weight of substance which can be removed in a series of extractions. Suppose we have a solution containing w grams of a substance in V_1 cc of solution, and suppose that this solution is shaken repeatedly with V_2 cc samples of pure immiscible second solvent until distribution equilibrium is attained. Then at the end of n extractions the weight w_n of solute remaining unextracted will be

$$w_n = w \left(\frac{KV_1}{KV_1 + V_2} \right)^n \tag{38}$$

and therefore the weight extracted will be

$$
\begin{aligned}
w - w_n &= w - w \left(\frac{KV_1}{KV_1 + V_2} \right)^n \\
&= w \left[1 - \left(\frac{KV_1}{KV_1 + V_2} \right)^n \right]
\end{aligned}
\tag{39}
$$

Here K is defined as $K = C_1/C_2$. When K is known for the distribution of the dissolved substance between any two given solvents, equation (38) may be employed to estimate the number of extractions necessary with V_2 cc of extracting solvent for V_1 cc of the original solution in order to reduce w to some given value w_n. Another important deduction which can be made from equation (38) is that if a given volume V of a solvent is available for extraction, greater extracting efficiency can be obtained if this volume is utilized in a number of separate extractions than if it were used in toto once. In other words, greater extracting efficiency is obtained by keeping V_2 small and n large than the other way around, and hence it is better to extract with small volumes of solvent several times than once with a large volume. The same conclusions apply to washing of precipitates, in which case the process may be considered as the distribution of the impurity between the wash liquid and the precipitate.

Another application of distribution coefficients is in analysis. Suppose a substance is present in a solvent A, in which analysis for the substance is difficult, and suppose, further, that analysis in another solvent B is readily possible. Then a distribution of the substance between the two

solvents can be carried out and the substance analyzed for in solvent B. From the result of analysis in B, the volumes of the two solvents used, and the distribution coefficient, K, for the substance between the two solvents, the weight of it present originally in A can be obtained.

REFERENCES FOR FURTHER READING

1. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.
2. K. Jellinek, "Lehrbuch der physikalischen Chemie," Ferdinand Enke, Stuttgart, 1930, Vol. III.
3. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931, Vol. I.

PROBLEMS

1. Formulate equilibrium constants for the following equilibria:
   - (a) $O_2(g) + 2\,Hg(l) = 2\,HgO(s)$
   - (b) $HCO_3^-(aq) + H^+(aq) = CO_2(g) + H_2O(l)$
   - (c) $NH_3(g) + H_2O(l) = NH_4OH(aq)$
   - (d) $MgO(s) + 2\,HCl(g) = MgCl_2(s) + H_2O(g)$
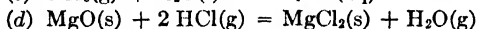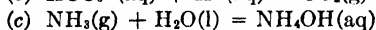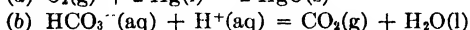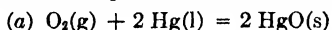
2. NH_4HS dissociates as follows:
$$NH_4HS(s) = NH_3(g) + H_2S(g)$$
At 25° C the dissociation pressure of the pure solid is 500 mm Hg. Calculate (a) K_p', and (b) the total pressure at equilibrium when 300 mm of NH_3 are introduced into a flask containing solid NH_4HS.

   <div align="right">*Ans.* (a) 6.25×10^4 mm²; (b) 141.5 mm</div>

3. For the equilibrium
$$C\ (graphite) + CO_2(g) = 2\,CO(g)$$
at 1123° K the mole per cent of CO in the vapor phase at equilibrium is 93.77% at 1 atm pressure. What is (a) K_p', and (b) the mole per cent of CO present at equilibrium when the total pressure is 10 atm?

4. In two experiments solid $NH_2CO_2NH_4$ is introduced at 30° C into a flask containing (a) a partial pressure of 200 mm of NH_3 gas, and (b) a partial pressure of 200 mm of CO_2 gas. Using the average value of K_p' in Table 3, find what will be the total pressure at equilibrium in each case.
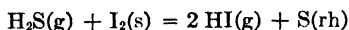
5. For the reaction
$$H_2S(g) + I_2(s) = 2\,HI(g) + S(rh)$$
$K_p' = 1.33 \times 10^{-5}$ atm at 60° C. What will be the mole fraction of HI in the vapor at 60° C when the total pressure is 1 atm?   *Ans.* 0.00365

6. At 713° K, K_p' for the reaction
$$Sb_2S_3(s) + 3\,H_2(g) = 2\,Sb(s) + 3\,H_2S(g)$$
is 0.429. What is the mole fraction of H_2 in the vapor at 713° K? Will the result depend upon the total pressure? Explain.
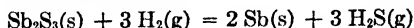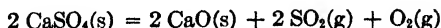
7. For the reaction,
$$2\,CaSO_4(s) = 2\,CaO(s) + 2\,SO_2(g) + O_2(g)$$
$K_p' = 1.45 \times 10^{-5}$ atm³ at 1625° K. What is the dissociation pressure of $CaSO_4$ in mm Hg at 1625° K?

8. One mole of H_2 and 1 mole of $Sb_2S_3(s)$ are introduced into a reaction vessel and heated to 713° K. From the data given in problem 6, find the number of moles of Sb formed, and the number of moles of H_2 unconverted.

9. One g of $Na_2HPO_4 \cdot 7\ H_2O$ is placed in a 2-liter vessel at 25° C. What weight of H_2O would have to be added in order to convert completely this hydrate into $Na_2HPO_4 \cdot 12\ H_2O$ and establish a condition of equilibrium in the vessel? Use the data given in Table 4.

10. For the transition

$$HgS\ (red)\ =\ HgS\ (black)$$
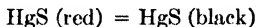
$\Delta F^0 = 4100 - 6.09\ T$. What is the stable modification of HgS at 100° C? What is the transition temperature?

11. For the reaction

$$2\ NaHSO_4(s)\ =\ Na_2S_2O_7(s)\ +\ H_2O(g)$$
$$\Delta H_{298.2°\ K} = 19{,}800\ cal \qquad \Delta F^0_{298.2°\ K} = 9000\ cal$$
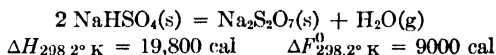
Assuming that ΔH is constant with temperature, calculate the dissociation pressure of $NaHSO_4$ at 700° K. *Ans.* 55.3 atm

12. For the reaction

$$CuSO_4 \cdot 3\ H_2O(s)\ =\ CuSO_4 \cdot H_2O(s)\ +\ 2\ H_2O(g)$$

the dissociation pressure at 298.2° K is 7.37×10^{-3} atm, while $\Delta H_{298.2°\ K} = 27{,}000$ cal. Assuming that ΔH does not vary with T, what is the dissociation pressure in mm Hg at 100° C?

13. The heat of solution of $MgCl_2(s)$ at 18° C is $-35{,}900$ cal, that of $MgCl_2 \cdot 6\ H_2O$ is -2950 cal, while the heat of vaporization of H_2O is 587 cal/gram at 18° C. Find ΔH at 18° C for the reaction
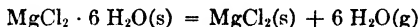
$$MgCl_2 \cdot 6\ H_2O(s)\ =\ MgCl_2(s)\ +\ 6\ H_2O(g)$$
Ans. 96,350 cal

14. At 20° C SO_2 was permitted to distribute itself between 200 cc of $CHCl_3$ and 75 cc of H_2O. When equilibrium was established, the $CHCl_3$ layer contained 0.14 mole of SO_2 and the H_2O layer 0.05 mole. What is the distribution coefficient of SO_2 between H_2O and $CHCl_3$ at 20° C? *Ans.* $C_{H_2O}/C_{CHCl_3} = 0.953$

15. Using the average value of the distribution coefficients given in Table 6, calculate the number of moles of H_3BO_3 which may be extracted from 50 cc of a 0.2 molar aqueous solution, (a) by a single extraction with 150 cc of amyl alcohol, and (b) by three extractions with 50-cc portions of amyl alcohol.

16. From the following data for the distribution of benzoic acid between H_2O and C_6H_6 at 20° C (a) show that benzoic acid is associated into double molecules in C_6H_6, and (b) calculate the distribution coefficients on the basis of this assumption. Neglect the dissociation of benzoic acid in water.

C_{H_2O}	$C_{C_6H_6}$
0.0150	0.242
0.0195	0.412
0.0289	0.970

17. Using the results of the preceding problem, calculate the number of moles of benzoic acid which may be extracted from 100 cc of a 0.2 molar aqueous solution by 10 cc of C_6H_6 at 20° C.

18. At 25° C the distribution coefficient of H_2S between H_2O and C_6H_6, defined as $[H_2S]_{H_2O}/[H_2S]_{C_6H_6}$, is 0.167. What is the minimum volume of C_6H_6 necessary at 25° C to extract in a single step 90% of the H_2S from 1 liter of a 0.1 molar aqueous solution of H_2S? *Ans.* 1.50 liters

19. Use the data of the preceding problem to find what total volume of C_6H_6 would be necessary to remove 90% of the H_2S from the given aqueous solution in three separate extractions using equal volume of C_6H_6 in each.

20. At 25° C the distribution coefficient of C_2H_5OH between CCl_4 and H_2O, $K = [C_2H_5OH]_{CCl_4}/[C_2H_5OH]_{H_2O}$, is 0.0244. How will 1 g of C_2H_5OH distribute itself between 20 cc of H_2O and 50 cc of CCl_4?

Heterogeneous Equilibrium II: The Phase Rule

Introduction. Up to this point various types of heterogeneous equilibria were considered from a number of different points of view. Thus, heterogeneous equilibria such as vaporization, sublimation, fusion, transition of one solid phase to another, solubility of solids, liquids, and gases in each other, vapor pressure of solutions, chemical reaction between solids or liquids and gases, and distribution of solutes between phases all have been approached by methods suitable for each particular type of equilibrium. These involved empirical rules, kinetic considerations, Raoult's law, Henry's law, the law of mass action, and the distribution law. However, it is possible to treat all heterogeneous equilibria from a unified standpoint by means of a principle called the *phase rule*. With this principle the number of variables to which each and every type of heterogeneous equilibrium is subject may be defined under various experimental conditions. By this definition the phase rule does not invalidate or supersede some of the methods of attack described for the quantitative study of such equilibria. The phase rule is merely able to fix the number of variables involved. The quantitative relations among the variables must be established through supplementary relations such as some of those mentioned above. The significance of this statement will become clearer as soon as the nature of the phase rule and the manner in which it is used are developed.

Definitions. Before proceeding to a statement of the phase rule, it will be necessary to define and explain in some detail certain terms which are employed frequently in this connection. These are system, phase, true, metastable, and unstable equilibrium, number of components, and degrees of freedom of a system.

A *system* is defined as a substance or mixture of substances isolated from all others in an inert container for the purpose of specific study of the effect of pressure, temperature, and the change of concentration on the state of the materials present. For example, the system "water" would consist of water placed alone in an inert container fitted with a movable piston and some heating or cooling device such that either the temperature of or the pressure on the system, or both of these variables,

may be varied at will. Similarly the system sodium chloride-water would consist of these two substances placed in such a container for purpose of studying the effect of the variables mentioned on the state of the mixture.

A *phase* is a homogeneous, physically distinct, and mechanically separable portion of a system. If in the system "water" ice, liquid water, and water vapor coexist, each form constitutes a separate phase. Each phase is separated from every other phase by a phase boundary. Furthermore, each phase may be continuous, such as a gas or liquid phase, or it may be broken up into a number of smaller portions, such as a group of ice crystals. The term "mechanically separable" in the definition means that each phase can be separated from every other phase by such operations as filtration, sedimentation, decantation, or by some other mechanical method of separation, say, hand picking of crystals. It does not include, however, such methods of separation as evaporation, distillation, adsorption, or extraction. Since all gases are completely miscible, only one gas phase is possible in a system. With liquids there appears to be no theoretical limit to the maximum number of phases possible, but six is the greatest number that have ever been observed in any one system. Apparently any number of solid phases is possible.

A state of *true equilibrium* is said to exist in a system when the same state can be realized by approach from either direction. Thermodynamically speaking, true equilibrium is attained when the free energy content of the system is at a minimum for the given values of the variables. An instance of such an equilibrium is ice and liquid water at 1 atmosphere pressure and 0° C. At the given pressure the temperature at which the two phases are in equilibrium is the same whether it is attained by partial melting of the ice or a partial freezing of the water. On the other hand, water at − 5° C can be obtained only by careful cooling of liquid water and not at all by fusion of ice at − 5° C. Water at − 5° C is said to be in a state of *metastable equilibrium*. Such a state can be realized only by careful approach from one direction, and may be preserved indefinitely provided the system is not subjected to sudden shock, stirring, or "seeding" by solid phase. As soon as a crystal of ice is introduced, solidification sets in rapidly, and the temperature rises to 0° C.

A state of *unstable equilibrium* is said to exist when the approach to equilibrium in a system is so slow that the system *appears* not to undergo any change with time. An instance of such a situation is sodium chloride dissolving into a solution which is very nearly saturated with the salt. Insufficient time of observation might make it appear that equilibrium had been reached, whereas actually the process is still proceeding very slowly toward true saturation. It must be realized that, although a metastable equilibrium represents a state of at least partial stability,

unstable equilibrium does not involve any equilibrium at all, but only a process of very slow change.

The *number of components* of a system is the *smallest number* of independently variable constituents, in terms of whose formulas equations may be written expressing the composition of each of the possible phases that may occur. The quantity desired here is the *smallest number*, and it is immaterial which particular constituents are chosen to express the compositions of the various phases. This point will become clearer from the following examples. In the system "water" the phases that occur are ice, liquid water, and water vapor. The composition of each of these phases can be expressed in terms of the single constituent water, and hence this is a one-component system. The variable could equally well be hydrogen or oxygen, for the specification of one of these automatically fixes the other through the formula H_2O. Similar considerations show that the minimum number of constituents necessary to describe the composition of all phases in the system sodium sulfate-water is two, and hence this is a two-component system. In this system the various phases that may occur are Na_2SO_4, $Na_2SO_4 \cdot 7\ H_2O$, $Na_2SO_4 \cdot 10\ H_2O$, solutions of Na_2SO_4 in water, ice, and water vapor. The composition of each of these phases in terms of the two components sodium sulfate and water may be stated as follows:

$$Na_2SO_4 : \quad Na_2SO_4 + 0\ H_2O$$
$$Na_2SO_4 \cdot 7\ H_2O : \quad Na_2SO_4 + 7\ H_2O$$
$$Na_2SO_4 \cdot 10\ H_2O : \quad Na_2SO_4 + 10\ H_2O$$
$$Na_2SO_4(aq.) : \quad Na_2SO_4 + x\ H_2O$$
$$H_2O(s),\ H_2O(l),\ H_2O(g) : \quad 0\ Na_2SO_4 + H_2O$$
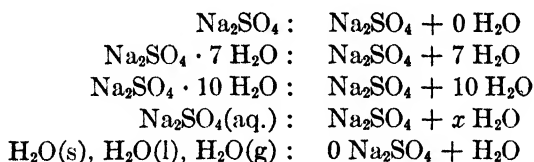
It will be noted that the composition of certain phases may be stated in terms of only one of these constituents, whereas certain others necessitate a knowledge of the amounts of both present in order to specify unambiguously the composition of the phase. Since two components are the smallest number by which the compositions of *all* the phases can be defined, sodium sulfate-water must be a two-component system.

The essential fact to remember in deciding upon the number of components of a system is that the particular constituents chosen as independent variables do not matter, but their *smallest* number does. If the number chosen is not the smallest, certain of these will not be independent of the others. Again, in writing the composition of a phase in terms of the components selected, plus, minus, and zero coefficients in front of a component are permissible. Thus, in a system in which solid magnesium carbonate dissociates according to

$$MgCO_3(s) = MgO(s) + CO_2(g) \tag{1}$$

the compositions of the various phases may be represented in terms of magnesium carbonate and magnesium oxide as follows:

$$MgCO_3: \quad MgCO_3 + 0\ MgO$$
$$MgO: \quad 0\ MgCO_3 + MgO$$
$$CO_2: \quad MgCO_3 - MgO$$
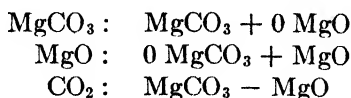
Finally, by the *degrees of freedom* or the *variance* of a system is meant the *smallest number* of independent variables (such as pressure, temperature, concentration) that must be specified in order to define completely the remaining variables of the system. The significance of the degrees of freedom of a system may be gathered from the following examples. In order to specify unambiguously the density of liquid water, it is necessary to state also the temperature and pressure to which this density corresponds; i.e., the density of water is 0.99973 gram per milliliter at 10° C and 1 atmosphere pressure. A statement of the density at 10° C without mention of pressure does not define clearly the state of the water, for at 10° C the water may exist at any and all possible pressures above its own vapor pressure. Similarly, mention of the pressure without the temperature leaves ambiguity. Hence, for complete description of the state of the water, two variables must be given, and this phase, when present alone in a system, possesses two degrees of freedom, or the system is said to be *bivariant*. When liquid and solid water exist in equilibrium, however, the temperature and the densities of the phases are determined only by the pressure, and a statement of some arbitrary value of the latter is sufficient to define all the other variables. Thus, if we know that ice and water are at equilibrium at 1 atmosphere pressure, the temperature can be only 0° C and the densities are also established. The same applies to the choice of temperature as the independent variable. At each arbitrarily chosen temperature (within the range of existence of the two phases) equilibrium is possible only at a given pressure, and once again the system is defined in terms of one variable. Under these conditions the system possesses only one degree of freedom, or it is *monovariant*.

The Gibbs Phase Rule. That there is a definite relation in a system between the number of degrees of freedom, the number of components, and the number of phases present was first established by J. Willard Gibbs in 1876. This relation, obtained by Gibbs through purely thermodynamic arguments, is known as the *phase rule*, and represents one of the most significant generalizations in the field of science. The phase rule is a principle of the widest generality, and its validity is in no way dependent on any concepts of atomic or molecular constitution. Although this rule has been known since 1876, it was not for some years that its tremendous importance was appreciated, and credit is due to

Ostwald, Roozeboom, van't Hoff, and others for showing with what effect this generalization can be utilized in the practical study of problems in heterogeneous equilibrium.

To arrive at a formulation of the phase rule, consider in general a system of C components in which P phases are present. The problem now is to determine the total number of variables or factors upon which such a system depends. First of all, the state of the system will depend upon the pressure and the temperature. Again, in order to define the composition of each phase, it is necessary to specify the concentration of $(C - 1)$ constituents of the phase: the concentration of the remaining component is then determined by difference. Since there are P phases, the total number of concentration variables will be $P(C - 1)$, and these along with the temperature and pressure constitute a total of $[P(C - 1) + 2]$ variables.

The student will recall from his study of algebra that when an equation in n independent variables occurs, n independent equations are necessary in order to solve for the value of each of these. Similarly, in order to define the $[P(C - 1) + 2]$ variables of a system, this number of equations relating these variables would have to be available. The next question is then: How many equations involving these variables can possibly be written from the conditions obtaining in the system? To answer this query recourse must be had to thermodynamics. Thermodynamics tells us that equilibrium between the various phases in a system is possible only provided the free energy [1] per mole of each constituent of a phase is equal to the free energy of the *same* constituent in every other phase. Since the free energy of the constituent of a phase is a function of the pressure, temperature, and $(C - 1)$ concentration variables, it readily follows that the thermodynamic condition for equilibrium makes it possible to write *one equation* among the variables *for each constituent distributed between any two phases*. When P phases are present, $(P - 1)$ equations are available for each constituent, and for C constituents a total of $P(C - 1)$ equations.

If this number of equations is equal to the number of variables, the system is completely defined. However, generally this will not be the case, and the number of variables will exceed the number of equations by F, where

$$F = \text{Number of variables} - \text{Number of equations}$$
$$= [P(C - 1) + 2] - [P(C - 1)]$$
$$= C - P + 2 \tag{2}$$

Equation (2) is the celebrated phase rule of Gibbs. F is the number of degrees of freedom of a system, and gives the number of variables whose values must be specified arbitrarily before the state of the system can

[1] Strictly speaking, the *partial* free energy.

be completely and unambiguously characterized. According to the phase rule the number of degrees of freedom of a system is determined by both the number of components and the number of phases present, or, rather, by $(C - P)$.

In this derivation it was assumed that each component is present in every phase. If a component is missing from a particular phase, however, the number of concentration variables is decreased by one. But at the same time the number of possible equations is also decreased by one. Hence the value of $(C - P)$, and therefore F, remains the same whether each constituent is present in every phase or not. This means that the phase rule is not restricted by the assumption made, and is generally valid under all conditions of distribution provided that equilibrium exists in the system.

Classification of Systems According to Phase Rule. The principal value of equation (2) is in its use as a check in the construction of various types of plots for the representation of the equilibrium conditions existing in heterogeneous systems. Before proceeding to a discussion of some specific systems and the application of the phase rule to these it is convenient to subdivide all systems into several major categories. The usual procedure is to classify all systems according to the number of components present. Thus we may have one-component, two-component, three-component systems, etc., depending on whether the system under consideration contains one, two, three, etc., independent components. The advisability of this method of approach will become apparent from what follows.

One-Component Systems. The complexity of one-component systems depends on the number of solid phases that can exist in the system. The simplest possible case is one in which only a single solid phase occurs, such as in water at the lower pressures, carbon dioxide, and a number of other systems. When more than a single solid phase appears in a system the number of possible equilibria is considerably enhanced, and hence the phase diagram, or the plot exhibiting the various possible equilibria, becomes more involved. The various possibilities in such systems and their phase relations can best be brought out by the consideration of several specific examples.

(a) **The System: Water.** Above about $-20°$ C and below 2000 atmospheres pressure there is only one solid phase in this system, namely, ordinary ice. This solid phase and liquid water and water vapor constitute the *three* possible single phases in the system. These three single phases may be involved in *three* possible two-phase equilibria, namely,

        (1) Liquid-vapor
        (2) Solid-vapor
        (3) Solid-liquid

and *one* three-phase equilibrium, solid-liquid-vapor. Applying the phase rule to the system, when only a single phase is present we see that $F = C - P + 2 = 1 - 1 + 2 = 2$, and therefore each single phase at equilibrium possesses two degrees of freedom. If temperature and pressure are chosen as the independent variables, the phase rule predicts that both of these must be stated in order to define the condition of the phase. But, since two independent variables are necessary to locate any point in an area, it must follow that each phase on a $P\text{-}T$ diagram occupies an area; and, as three single phases are possible in this system, we may anticipate three such areas on the plot, one for each phase.

When two phases coexist at equilibrium, the phase rule predicts that $F = 1 - 2 + 2 = 1$, or one degree of freedom. Since a single variable determines a line, we may expect for each two-phase equilibrium a line on the $P\text{-}T$ plot. As three such equilibria may occur in the system, the diagram will be characterized by the existence of three lines separating the various areas from each other. Finally, for the three-phase equilibrium $F = 1 - 3 + 2 = 0$, i.e., no variables need be specified in order to define this equilibrium. This must mean that when all three phases coexist the temperature and pressure are fixed, and the position of this equilibrium on the diagram is characterized by the intersection of the three lines at a common *point*.

Although the phase rule makes it possible to predict in the manner described the general appearance of the diagram, the exact positions of all lines and points can be determined only by experiment. An inspection of the possible equilibria in this system shows that the data necessary for the construction of a $P\text{-}T$ plot are: (1) the vapor pressure curve of water (liquid-vapor equilibrium), (2) the sublimation curve of ice (solid-vapor equilibrium), (3) the melting point curve of ice as a function of pressure (solid-liquid equilibrium), and (4) the position of the solid-liquid-vapor equilibrium point, or triple point. These experimental data for the system water are shown in
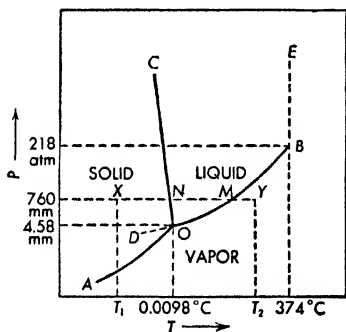
Fig. 1. The System Water at Moderate Pressures (Schematic)

Fig. 1. In this phase diagram line AO gives the sublimation curve of ice, line OB the vapor pressure curve of liquid water, and OC the line along which equilibria between ice and liquid water occur at various pressures. O is the *triple point* at which ice, water, and water vapor are at equilibrium. This equilibrium is possible only at 0.0098° C and 4.58 mm pressure. As predicted by the phase rule, there are three areas on the

diagram, disposed as shown, one such area corresponding to each of the single phases.

The vapor pressure curve of water, line OB, extends from the triple point O up to the critical point B, corresponding to 374° C and 217.7 atmospheres pressure. However, under certain conditions it is possible to supercool water below point O to yield the metastable liquid-vapor equilibria shown by the dotted line OD. The fact that OD lies above AO shows that at temperatures below that of the triple point liquid water has a vapor pressure higher than the sublimation pressure of ice, and hence the supercooled liquid is unstable at these temperatures with respect to the ice. The sublimation curve of ice, line AO, may extend from absolute zero up to O. No superheating of ice beyond O has ever been realized. Line CO runs from O up to a point corresponding to 2000 atmospheres pressure and about − 20° C; at this point ordinary ice, type I, in equilibrium with water undergoes a transformation into another solid modification, type III, in equilibrium with the liquid. The slope of this line indicates that the melting point of ice is *lowered* by increase in pressure in accord with the Le Chatelier principle and the fact that ice has a larger specific volume than liquid water. The slopes of the lines AO, OB, and OC are determined at each point by the Clapeyron equation or one of its modifications as applicable in each instance. From the slopes and this equation it is possible to evaluate the heats of vaporization from OB, the heats of sublimation from AO, and the heats of fusion from OC.

Since no liquid may exist above the critical temperature, dotted line BE has been inserted in the diagram to separate the liquid from vapor areas above the critical temperature. Consequently the vapor area lies below and to the right of $AOBE$, the liquid area above OB and between the lines OC and BE, while the solid area extends to the left of OC and above AO.

The manner in which a diagram such as Fig. 1 may be used to follow the changes that occur in the system with a change in the variables may be seen from the following example. Suppose it is desired to know the behavior of the system on heating ice at a pressure of 760 mm and temperature T_1, corresponding to point X in the diagram, to a point corresponding to point Y at the same pressure but temperature T_2. Starting with ice at X and heating it slowly at constant pressure, the system follows line XN with an increase in the temperature of the ice and a slight decrease in its density. However, once N is reached the ice begins to melt, the temperature remains constant until the fusion is complete, and only when this process is terminated does the temperature begin to rise again along NM. Between N and M the only change is an increase in the temperature of the liquid and a decrease in its density. But at M vaporization sets in, and the temperature again is constant until all the

liquid is converted to vapor. On complete transformation of liquid to vapor, any further addition of heat results merely in an increase of the temperature of the vapor along MY until the latter point is reached. In the same way it is possible with the aid of Fig. 1 to predict and outline any changes that may take place in this system with a variation of temperature, pressure, or both.

At high pressures a number of other solid modifications besides the ordinary form have been observed. The phase diagram of the system under these conditions is shown in Fig. 2. Figures 1 and 2 are parts of the same diagram, the latter figure being merely the high-pressure portion of the former. The manner in which the two diagrams tie in can be judged from the line OC, which is the same in both plots. It is of interest to observe in Fig. 2 that at very high pressures ice of types VI and VII may exist at temperatures above 0° C. In fact, at a pressure of about 40,000 atmospheres ice VII is stable at 190° C! For further details on the relations in Fig. 2 the student must be referred to the literature.[1]

Fig. 2. The System Water at High Pressures

(b) *The System: Sulfur.* Sulfur exists in two solid modifications, the *rhombic* form stable at ordinary temperatures, and the *monoclinic* variety stable at higher temperatures. These two solid phases along with the liquid and vapor give a possibility of existence of four single phases, which in turn can lead to the following equilibria:

Two-Phase Equilibria	Three-Phase Equilibria
(1) S(r) — S(vapor)	(1) S(r) — S(m) — S(liquid)
(2) S(m) — S(vapor)	(2) S(r) — S(liquid) — S(vapor)
(3) S(r) — S(liquid)	(3) S(m) — S(liquid) — S(vapor)
(4) S(m) — S(liquid)	(4) S(r) — S(m) — S(vapor)
(5) S(liquid) — S(vapor)	Four-Phase Equilibria
(6) S(r) — S(m)	(1) S(r) — S(m) — S(liquid) — S(vapor)

Applying the phase rule to these possible equilibria, we may anticipate four separate divariant single-phase areas, six monovariant two-phase equilibrium lines, and four invariant three-phase equilibrium points.

[1] Tammann, Zeit. physik. Chem., **72**, 609 (1910); Bridgman, Proc. Am. Acad. Sci., **47**, 441 (1912); J. Chem. Phys., **5**, 964 (1937).

Since the maximum number of phases that may be present together in equilibrium is given by the phase rule with $F = 0$, it follows that $P = C + 2 - F = 3$ for a one-component system, and hence the four-phase equilibrium cannot exist in this system or in any other one-component system.

Figure 3 shows the experimentally determined phase diagram for this system. The four single-phase areas are disposed as indicated. Lines OP and PT are the sublimation curves of rhombic and monoclinic sulfur, respectively, while TU is the vapor pressure curve of liquid sulfur. At point P rhombic sulfur undergoes a transition to monoclinic sulfur, and hence this is an invariant point corresponding to the equilibrium $S(r) - S(m) - S(vapor)$. At point T monoclinic sulfur melts, and thus this point corresponds to the three-phase equilibrium $S(m) - S(liquid) - S(vapor)$. Line PS shows the variation of the transition point with

Fig. 3. The System Sulfur (Schematic)

pressure, while line TS shows in a like manner the variation of the melting point of monoclinic sulfur with the same variable. These two lines intersect at S to yield the three-phase equilibrium $S(r) - S(m) - S(liquid)$. Finally, the line SW gives the melting point of rhombic sulfur. These are all the *stable* equilibria that occur. The monoclinic sulfur area is enclosed by the lines PS, PT, and TS, and therefore no monoclinic sulfur can possibly exist in stable condition outside this area. Furthermore, no vapor can exist stably at pressures above those given by the lines OP, PT, and TU below the temperature of the critical point U. Above this temperature, however, no liquid is possible. Hence the liquid area is terminated along the vertical dotted line through U, and vapor may exist thereafter at high pressures.

The remaining equilibria in this system are all *metastable*. By rapid heating it is possible to superheat rhombic sulfur along the extension of OP, the line PR, which is the equilibrium line for superheated $S(r)$ with $S(vapor)$. Similarly it is possible to supercool liquid sulfur along the extension of TU to R. At point R the metastable $S(r) - S(vapor)$ line intersects the metastable $S(liquid) - S(vapor)$ line to yield the metastable invariant triple point $S(r) - S(liquid) - S(vapor)$. The line RS shows the variation of this metastable point with pressure, and is, therefore, the melting point line for rhombic sulfur as a function of pressure. This line is an extension of the stable SW line into the metastable range.

It should be clearly understood that, when these various metastable equilibria occur, monoclinic sulfur does not appear; instead rhombic sulfur is transformed directly to liquid along RS or to vapor along PR without passing through the monoclinic stage.

In order to clarify the relations in this system, consider conversion of rhombic sulfur at a pressure P_1 and temperature T_1, point X, to vapor at the same pressure but temperature T_2, point Y. If the rhombic sulfur is heated slowly at constant pressure, its temperature will rise until point Z is reached. At this point, allowing sufficient time for establishment of equilibrium, the rhombic sulfur will transform at constant temperature to the monoclinic form. When this transition is over, the temperature will again climb to N, the melting point, where the monoclinic sulfur is converted to liquid at the constant temperature corresponding to N. Once this process is completed, further addition of heat will result in the heating of the liquid to Q, where vaporization takes place under isothermal conditions, and thereafter the temperature of the vapor will rise steadily to Y. Throughout all these operations only stable equilibria were encountered. If, however, the rhombic sulfur is heated rapidly, it may be carried past Z to M without transformation. At M the rhombic sulfur melts to liquid which subsequently vaporizes at Q. By operating in this manner with the metastable states between Z and N no monoclinic form is encountered, and the rhombic sulfur goes directly to the liquid.

(c) *Other One-Component Systems.* The system *carbon dioxide* is very similar to the system water as shown in Fig. 1, except that the triple point occurs at $-56.4°$ C and at a pressure of about 5 atmospheres. Furthermore, the solid-liquid line slopes to the right instead of to the left as in the water diagram. Since the one atmosphere line in this system cuts only the sublimation curve of the solid, solid carbon dioxide must change directly to vapor at this pressure without passing through the liquid state. Liquefaction of the solid can be attained only under pressures of about 5 atmospheres or higher.

Other examples of one-component systems investigated are phosphorus and benzophenone. A detailed discussion of these is given by Findlay and Campbell.[1]

2) Two-Component Systems. When a single phase is present in a two-component system, the number of degrees of freedom is $F = 2 - 1 + 2 = 3$. This means that three variables must be specified in order to describe the condition of the phase: pressure, temperature, and the concentration of one of the components. In order to present these relations graphically, three coordinate axes at right angles to each other

[1] Findlay and Campbell, "The Phase Rule," eighth edition, Longmans, Green and Company, New York, 1938.

would be required, and the diagram thus resulting would be a solid figure. Since such figures are difficult to construct and use, the more common practice is to employ either a projection of such a solid diagram on a plane, or a planar cross section of the figure for a given constant value of one of the variables. In this manner it is possible to present the various relations in two-component systems in a two-dimensional plot of any two of the three variables mentioned.

Moreover, the discussion of two-component systems can be simplified further by considering the various possible types of equilibria separately rather than attempting to present them all at once. Thus the usual practice is to study liquid-gas, solid-gas, liquid-liquid, and solid-liquid equilibria individually, and, when necessary, to combine the diagrams. Since the first three types of equilibria have already been considered in various places in the text, attention here will be devoted exclusively to an application of the phase rule to the very important category of solid-liquid equilibria.

Solid-Liquid Equilibria. Solid-liquid equilibria are of great importance because of their connection with all crystallization problems. Such equilibria are characterized generally by the absence of a gas phase, and by the fact that they are little affected by small changes in pressure. Systems where the gas phase is absent are called *condensed systems*, and it is with condensed systems that we shall be concerned.

Measurements on solid-liquid equilibria in condensed systems are generally carried out at atmospheric pressure. Because of the relative insensitivity of such systems to small variations in pressure, the latter may be considered constant, and for such systems the phase rule takes the form

$$F = C - P + 1 \tag{3}$$

For two-component systems equation (3) becomes

$$F = 2 + 1 - P$$
$$= 3 - P \tag{4}$$

where the only remaining variables are temperature and the concentration of one of the constituents. Solid-liquid equilibria are represented, therefore, on temperature-composition diagrams. For limited ranges of concentration any scheme of expressing concentration will do, but where the range may extend from 100 per cent of one constituent to 100 per cent of the other, it is preferable to employ as an abscissa either weight percentage or mol percentage, as in distillation diagrams.

Experimental Methods for Solid-Liquid Equilibria. Many experimental procedures have been employed from time to time for the determination of equilibrium conditions between solid and liquid phases. Of these the two of widest utility and applicability are the *thermal anal-*

ysis and *saturation or solubility methods.* These methods, supplemented when necessary by an investigation of the nature of the solid phases occurring in a system, can cover between them the study of any system which may be encountered.

(a) **The Thermal Analysis Method.** The thermal analysis method involves a study of the cooling rates, i.e., temperature-time curves, of the various mixtures of a system during solidification. From such curves it is possible to ascertain the temperature of initial and final solidification of a mixture, and to detect the temperatures at which various transformations and transitions occur. Although thermal analysis is applicable under all temperature conditions, it is particularly suitable for equilibrium investigations at temperatures considerably above and below that of the room.

In order to illustrate the experimental steps involved in this procedure, the interpretation of the curves, and the plotting of the final diagram, consider specifically the problem of determining the condensed phase diagram for the binary system bismuth-cadmium. The first step involves the preparation of a number of mixtures of the two metals ranging in overall composition from 100 per cent bismuth to 100 per cent cadmium. These mixtures may be spaced at 10 per cent intervals, and should all be preferably of equal weight, say 100 to 300 grams. Each of these mixtures of solids is placed in an inert crucible of, say, fireclay or graphite and is then melted in an electric furnace. To prevent oxidation of the metals it is advisable to maintain an inert or reducing atmosphere over them by passing hydrogen, nitrogen, or carbon dioxide through the furnace. A molten flux, such as borax, or a layer of powdered graphite may be used to cover the crucible charge as an added precaution. After melting and thorough agitation to obtain a uniform molten mass, a thermocouple is inserted in the melt, and the furnace and contents are allowed to cool slowly. Temperature and time readings are taken at definite time intervals, usually every 15 to 30 sec, until the charge in the crucible is completely solidified. For taking temperature-time readings recording potentiometers are particularly convenient, for they yield a continuous and permanent record of temperatures at various times during the cooling. Finally, plots of the temperatures thus obtained against time are prepared. If a second check on the composition is desired, the solidified alloys are removed and carefully analyzed.

Figure 4 shows a set of cooling curves thus obtained for various compositions of bismuth-cadmium mixtures. The explanation of these curves is as follows. When a body that liberates only sensible heat is cooled slowly and uniformly, a smooth cooling curve is obtained, and the temperature of the body approaches that of the room as a limit. However, when some transformation that liberates heat occurs during cooling, the

slope of the cooling curve is reduced suddenly. The nature of the reduction depends on the degrees of freedom of the system. A single phase with $F = 2$ exhibits a continuous cooling curve. When a new phase appears, the variance of the system is reduced to one, and the heat generated by the formation of the new phase results in a discontinuity in the curve due to change of slope for the cooling of one phase to a lesser slope
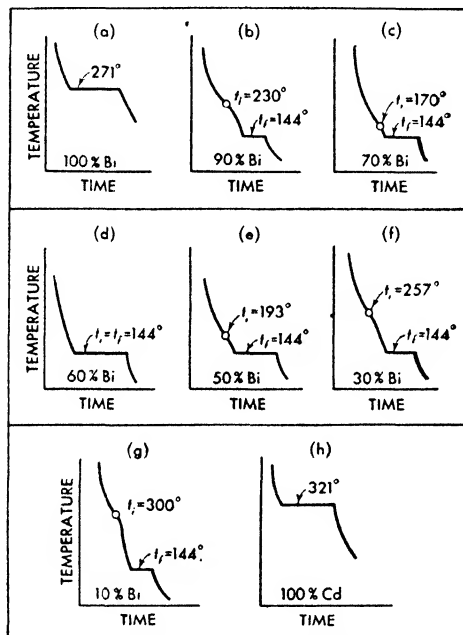
Fig. 4. Cooling Curves for the System Bismuth-Cadmium

corresponding to the cooling of two phases. Again, when still a third phase appears, $F = 0$, and the temperature of the system must remain constant until one of the phases disappears. The result is a flat portion on the cooling curve. Finally, when solidification is complete, the system regains some degree of freedom, and the cooling curves once again exhibit continuous variation of temperature with time. In light of these arguments a "break" or arrest in a cooling curve indicates the appearance of a second phase, usually the separation of a solid from the melt, while a horizontal portion indicates the coexistence of three phases. The third phase may result from the separation of two solids from the melt, the interaction of the melt and one solid to form another solid, or the separation of a solid from two liquid phases. The nature of the particular change occurring can be ascertained from an inspection of the final phase diagram and an analysis of the solids in the system.

With these considerations in mind we may conclude from the cooling curves in Fig. 4 that the arrests indicated by t_i signify the appearance of a second phase in the system, while the horizontal portions result from the coexistence of three phases. In this system the only solid phases are pure bismuth and pure cadmium, and hence the horizontal portions are the result of simultaneous occurrence of these solids and melt. However, in curves a and h the horizontal portions are due to two phases, since these are one-component systems.

To construct the equilibrium diagram from the cooling curves, the initial and final solidification temperatures, t_i and t_f, are taken off the cooling curves for the various overall concentrations, and are plotted on a temperature-composition diagram. Smooth curves are drawn then through all the t_i and t_f temperatures to yield the diagram shown in Fig. 5. Curve AB indicates the temperatures at which bismuth begins to separate from various concentrations of melt, while BC gives the same information for initial separation of cadmium. Line DE indicates the temperature at which all mixtures become completely solid. Further details of this type of diagram will be discussed later.

Fig. 5. Phase Diagram for the System Bismuth-Cadmium

(b) *The Saturation or Solubility Method.* In Fig. 5 line AB may be looked upon not only as the initial freezing point curve for bismuth from bismuth-cadmium melts, but also as the *solubility* curve of bismuth in molten cadmium. Points on this curve represent then the solubilities of bismuth in the molten cadmium at various temperatures. Similarly curve BC gives the solubilities at various temperatures of cadmium in molten bismuth, while at B the solution is saturated with respect to both solids.

In the saturation method the solubilities of one substance in another are determined at various constant temperatures, and the solubilities are then plotted as a function of the temperature. To obtain the composition of a solution of cadmium saturated with bismuth at say 200° C, point F in Fig. 5, excess powdered bismuth may be added to molten cadmium, the mixture brought to 200°, and the mass agitated until equilibrium is attained. The excess solid bismuth is now filtered off, and the saturated solution is analyzed for both constituents. By repeating this operation at various temperatures between 144 and 271° C, curve AB may be traced out. By a similar procedure, but employing molten bismuth and excess solid cadmium, curve BC may be obtained between 144

and 321° C. Although this method is rarely applied for the study of metallic systems, it is the principal means employed in systems containing water and similar solvents. Outside temperatures between $-50°$ C and 200° C this method is attended by many experimental difficulties, and the thermal analytical procedure is consequently preferred.

Determination of Nature of Solid Phases. For a complete interpretation of the phase diagram it is essential to know the nature and composition of the solid phases which appear during crystallization and in the final solid. These may be:

(1) *Pure components*, such as bismuth or cadmium in the system discussed above.

(2) *Compounds* formed by reaction between the pure constituents. Examples are $MgZn_2$ in the system Mg-Zn, $Na_2SO_4 \cdot 10 H_2O$ in the system $Na_2SO_4-H_2O$, and $FeCl_3 \cdot 6 H_2O$ in the system $FeCl_3-H_2O$. Such compounds have definite composition, are stable over definite temperature and solution concentration intervals, possess definite melting or transition temperature, and have a definite crystal structure.

(3) *Solid solutions*, homogeneous solids whose composition may vary within certain concentration intervals, and is determined by the composition of the solution from which the solid crystallizes. Such variability in composition differentiates the solid solution from a compound.

(4) *Mixtures* of solids which may be pure components, compounds, or solid solutions.

Frequently it is possible to deduce the nature of the solid phases from the shape of the phase diagram, as will be seen. However, in some instances it is necessary to subject the solids to a more careful scrutiny. To do this the completely solidified mass may be inspected under a microscope to determine the number of solids present and, if possible, their identity. Another possibility is the use of x rays. Direct separation of solids from a solution and their subsequent analysis may also be resorted to, but this procedure is attended by some uncertainty because of difficulty in obtaining a solid free from contamination by saturated solution. This difficulty may be circumvented sometimes by the use of a "telltale." The "telltale" is a substance that is soluble in the solution but not in the solid phase. On adding a definite small quantity of this substance to the solution it is possible, by analyzing the amount of this substance present in the wet solid, to ascertain the quantity of solution adhering to the solid and thereby arrive at the composition of the pure solid phase.

Classification of Two-Component Solid-Liquid Equilibria. Every condensed phase diagram may be considered to be composed of a combination of a number of simple types of diagrams. In some systems only a single type occurs; in others a number of these simple types

may occur combined to yield a more complicated complete diagram. In either case the phase relations in a system can readily be understood when the significance of the elementary types of diagrams is manifest. For this reason we proceed to a classification and discussion of various simple types of diagrams which may be encountered in condensed two-component systems.

All condensed two-component systems may be classified first according to the miscibility of the liquid phases, and these in turn according to the nature of the solid phases which crystallize from the solution. On this basis the important elementary types are:

Class A.   The two components are completely miscible in the liquid state.
        Type I.   The pure components only crystallize from the solution.
        Type II.   The two constituents form a solid compound stable up to its melting point.
        Type III.   The two components form a solid compound which decomposes before attaining its melting point.
        Type IV.   The two constituents are completely miscible in the solid state and yield thereby a complete series of solid solutions.
        Type V.   The two constituents are partially miscible in the solid state and form stable solid solutions.
        Type VI.   The two constituents form solid solutions which are stable only up to a transition temperature.
Class B.   The two components are partially miscible in the liquid state.
        Type I.   Pure components only crystallize from solution.
Class C.   The two components are immiscible in the liquid state.
        Type I.   Pure components only crystallize from solution.

Class A: Type I. Simple Eutectic Diagram. Two-component condensed systems belonging to this class have a diagram of the general form shown in Fig. 6. They are characterized by the fact that two constituents A and B are completely miscible in the liquid state, and such solutions yield only pure A or pure B as solid phases. In this figure points D and E are the melting points of pure A and pure B respectively. Line DG gives the concentrations of solutions saturated with A at temperatures between D and F, or the freezing points of the solutions which yield A as a solid phase. Similarly line EG gives the concentrations of solutions saturated with solid B at temperatures between E and F. At G the solution is saturated with both A and B, i.e., at G three phases are in equilibrium. It follows, therefore, that the lines DG and EG represent monovariant two-phase equilibria, while G is an invariant point. At this point the temperature F and the composition C of the solution must remain constant as long as three phases coexist. The temperature can be lowered below F only when one of the phases has disappeared,

and on cooling this must be the saturated solution. In other words, at F solution G must solidify completely. Temperature F must consequently be the *lowest* temperature at which a liquid phase may exist in the system A–B; below this temperature the system is completely solid. The temperature F is called the *eutectic temperature*, the composition C the *eutectic composition*, point G the *eutectic point* of the system.

The area above the lines DG and GE is the area in which unsaturated solution or melt exists. In this area only one phase is present and the system is divariant. In order to define any point in this area both the temperature and composition must be specified. The significance of the remaining portions of the diagram can be made clear by considering the behavior on cooling of several mixtures of A and B. Take first a mixture of overall composition a. If such a mixture is heated to point a''' an unsaturated solution is obtained. On cooling this solution nothing beyond a drop of temperature of the liquid phase occurs until point a'', corresponding to temperature x'', is reached. At this point the solution becomes saturated with A; or, in other words, a'' is the freezing point of the solution at temperature x''. As cooling continues A keeps on separating out, and the composition of the saturated solution changes along the line $a''G$. Thus at a temperature such as x' solid A is in equilibrium with saturated solution of composition y', and so on. It is seen, therefore, that for any overall composition falling in area DFG solid A is in equilibrium with various compositions of solution given by the curve DG at each temperature. However, at temperature F another solid phase, B, appears, and the system becomes invariant. On extracting heat from this system A and B must crystallize from the saturation solution in the fixed ratio C, and this crystallization will proceed until the solution has been completely solidified. Once this process is over and nothing but a *mixture* of solid A and solid B remains, the system becomes monovariant, and the cooling may be continued below F into area $FACG$, the area of coexistence of the two solids A and B.

An inspection of the solid in area $FACG$ under a microscope would reveal that it is composed of relatively large crystals of A which have had a chance to grow from temperature x'' to the temperature below F, and an intimate *mixture* of finer crystals of A and B which crystallized in definite proportion C at temperature F. The larger crystals of A are

Fig. 6. Simple Eutectic Diagram

called *primary* crystals, since they appeared first. Area $FACG$ may be marked, therefore, as containing primary crystals of A and an intimate eutectic mixture of fine crystals of A and B.

Similar considerations applied to overall compositions lying between C and B, such as b, for instance, show that in area EFG solid B is in equilibrium with saturated solutions along the line EG. At temperature F solid A also appears. The system becomes invariant and remains so until the solution at G solidifies in the manner described. Once the solidification is complete, the mixture passes into area $FBCG$ where now primary B and eutectic mixture of overall composition C are present. Finally, an inspection of the diagram reveals that cooling a mixture of overall composition C will yield no solid until point G is reached. At this point solids A and B appear simultaneously, and the system solidifies at constant temperature to yield only the eutectic mixture. In this respect a mixture such as C behaves like a pure substance on freezing. However, the result is not a single solid, but a mixture of two.

Once a phase diagram such as Fig. 6 is available for a binary system, it is readily possible to specify the conditions under which particular solid phases may be obtained, and to describe in detail the behavior of any given overall mixture on cooling. Thus it is seen that pure A may be separated only from mixtures falling in area DFG, and only between the temperatures D and F. Similarly pure B may be obtained only in area EFG from overall compositions between C and B, and only between the temperatures E and F. The proportion of solid to saturated solution at each temperature can be estimated graphically from the diagram. For overall composition a at a temperature such as x', the distance $x'a'$ is a measure of the amount of saturated solution of composition y', while the distance $a'y'$ is a measure of the amount of solid A present in the mixture. Consequently the ratio $x'a'/a'y'$ is the weight ratio of y' to A if the composition is expressed in weight percentage, or of the number of moles of y' to A if the composition is expressed in mol percentage. From these ratios and any overall weight it is possible to calculate the yield of solid phase to be anticipated at any given temperature.

Examples of systems exhibiting simple eutectic diagrams of the type shown in Fig. 6 are aluminum-tin, bismuth-cadmium, potassium chloride-silver chloride, sodium sulfate-sodium chloride, and benzene-methyl chloride. The behavior of all these systems on cooling will be similar to that described for Fig. 6 and the system bismuth-cadmium.

Class A: Type II. Formation of Compound with Congruent Melting Point. When the two pure components react to form a compound stable up to its melting point, the phase diagram takes on the typical form shown in Fig. 7 for the system cuprous chloride-ferric chloride. If the compound, in this case $CuCl \cdot FeCl_3$, is considered as a separate

component, the whole diagram may be thought of as being composed of two diagrams of the simple eutectic type, one for $CuCl$-$CuCl \cdot FeCl_3$, and the other for $CuCl \cdot FeCl_3$-$FeCl_3$. The discussion of type I diagrams can now be applied to each portion of the diagram with the results shown in the figure.

A compound such as $CuCl \cdot FeCl_3$ which melts at a constant temperature to yield a liquid of the same composition as the solid compound is said to have a *congruent* melting point. Compounds

Fig. 7. Formation of Compound with Congruent Melting Point

with congruent melting points are also formed in such binary systems as gold-tellurium ($AuTe_2$), aluminum-selenium (Al_2Se_3), calcium chloride-potassium chloride ($CaCl_2 \cdot KCl$), urea-phenol (1 : 1), and many others. When several compounds with congruent melting points are formed in a
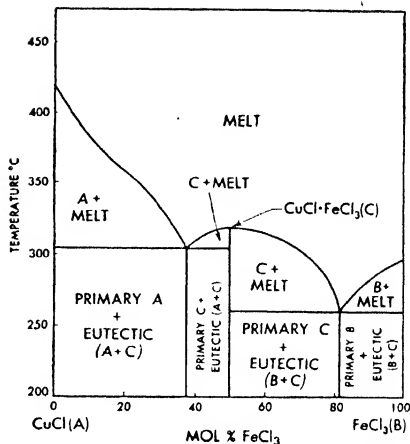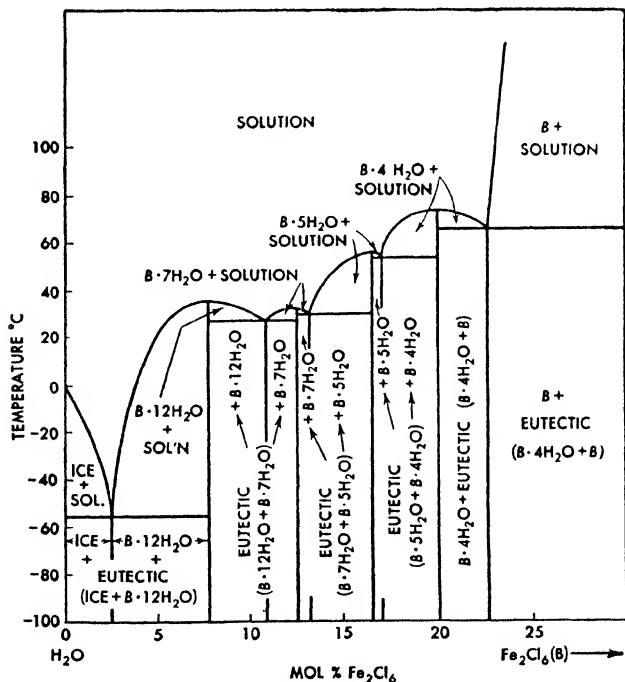
Fig. 8. The System Fe_2Cl_6-H_2O

system, a maximum is obtained for each compound with the result that the diagram takes on the appearance of Fig. 8 for the system Fe_2Cl_6-H_2O.[1] In this system four stable compounds have been discovered, namely, $Fe_2Cl_6 \cdot 12 H_2O$, $Fe_2Cl_6 \cdot 7 H_2O$, $Fe_2Cl_6 \cdot 5 H_2O$, and $Fe_2Cl_6 \cdot 4 H_2O$. The occurrence of these compounds increases the number of various areas and eutectic points, but otherwise introduces nothing new. By employing the artifice used for Fig. 7 of considering each compound as a constituent, the significance of the various areas, lines, and points can readily be deduced to be as indicated.

Class A: Type III. Compound Formation as Result of Peritectic Reaction. In many systems compounds are formed whose stability does not extend all the way to the melting point. When such compounds are heated, it is found that instead of melting congruently they decompose to yield a new solid phase and a solution composition *different* from that of the solid phases. When this happens the compound is said to undergo a *transition*, or *peritectic reaction*, or *incongruent fusion*.

Any peritectic reaction can in general be represented by the equation

$$C_2 \rightleftharpoons C_1 + \text{solution (or melt)} \tag{5}$$
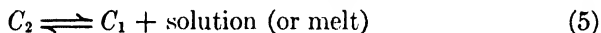
where C_2 is the compound, and C_1 is the new solid phase which may be itself a compound or the pure constituent. As equation (5) indicates, the peritectic reaction is reversible, i.e., on heating the change from left to right will occur, while on cooling the reverse will take place. Since during the peritectic reaction three phases are present at equilibrium, the system is invariant, and hence the temperature as well as the compositions of all the phases are fixed as long as the three phases are present. Temperature or composition can change only when one of the phases disappears, or, in other words, when the peritectic reaction is completed. The constant temperature at which the peritectic reaction takes place is called the *peritectic* or *transition temperature*, and for this temperature there will be obtained a horizontal portion on the cooling curves analogous to the eutectic portion.

The relations in a system in which a peritectic reaction occurs can best be illustrated with an actual example. For this purpose consider the diagram for the binary condensed system calcium fluoride-calcium chloride shown in Fig. 9. Line AB gives the concentrations of melt in equilibrium with solid calcium fluoride. When point B, corresponding to 737° C, is reached, a peritectic reaction between the melt of composition B and solid calcium fluoride sets in to form the double compound $CaF_2 \cdot CaCl_2$ according to the reaction

$$CaF_2(s) + \text{melt } (B) = CaF_2 \cdot CaCl_2(s) \tag{6}$$
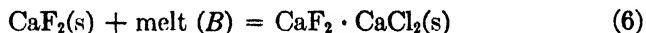
[1] Roozeboom, Z. physik. Chem., **10**, 477 (1892).

This reaction proceeds isothermally at 737° C to form the compound of composition C until either all the calcium fluoride or all of the melt has been consumed. Whether one or the other of these phases disappears depends on the overall composition of the mixture. If the overall com-

position lies between calcium fluoride and the compound C, such as a in the figure, there is more calcium fluoride present than is necessary to react with all of the melt B, and hence on completion of the reaction all the melt will be converted to $CaF_2 \cdot CaCl_2$, while the excess calcium fluoride will remain as such. Consequently, on passing into the area between the calcium fluoride axis and C below 737° C, only the two solid phases, CaF_2 and $CaF_2 \cdot CaCl_2$, will be present. If, however, the overall composition of the mixture lies between C and B, such as d, more melt is
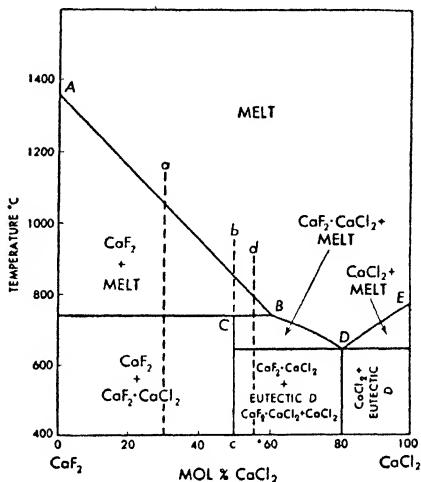
Fig. 9. Phase Diagram for the System CaF_2-$CaCl_2$

present than is required to react with all of the solid calcium fluoride, and hence the products of the peritectic reaction will be C and unreacted melt. On passing into the area immediately below CB we shall have, therefore, solid C and melt in equilibrium with the concentrations of the saturated solutions given by the line BD. It should be observed that the lines AB and BD are not continuous, but that there is a break at B. This means that the solid phases calcium fluoride and $CaF_2 \cdot CaCl_2$ each have separate and distinct solubility curves which intersect at B, the concentration of solution saturated with both phases. Finally, should the overall composition happen to be b, corresponding exactly to C, there is just sufficient calcium fluoride to react with all of the melt, and the result of the peritectic reaction is now only pure compound C.

The remainder of the diagram introduces nothing new. $CBDE$ is merely a simple eutectic type diagram involving the constituents C and calcium chloride. From the interpretation of such a diagram the area designations shown in the figure readily follow. It will be observed that the eutectic mixture is composed of calcium chloride, $CaF_2 \cdot CaCl_2$, and melt, with no pure calcium fluoride appearing.

Peritectic reactions have been observed in many binary systems. As examples may be quoted the following, along with the compounds formed: gold-antimony ($AuSb_2$), potassium chloride-cupric chloride

(2 KCl · CuCl$_2$), picric acid (A)-benzene (B) ($A \cdot B$), sodium chloride-water (NaCl · 2 H$_2$O), and sodium sulfate-water (Na$_2$SO$_4$ · 10 H$_2$O). In some systems several compounds occur, some of which have congruent melting points while others do not. Thus in the system magnesium-nickel the compound MgNi$_2$ has a congruent melting point, while the compound MgNi is formed as a result of a peritectic reaction. Similarly in the system aluminum-calcium the compound Al$_2$Ca has a congruent melting point, while the compound Al$_3$Ca does not. The phase diagram for the latter system, shown in Fig. 10, may be considered typical of

Fig. 10. Phase Diagram for the System Aluminum-Calcium

the relations encountered under such conditions. Finally, in many systems several compounds are formed none of which has a congruent melting point. Instances of this type are potassium sulfate-cadmium sulfate (K$_2$SO$_4$ · 2 CdSO$_4$, K$_2$SO$_4$ · 3 CdSO$_4$) and magnesium sulfate-water (MgSO$_4$ · H$_2$O, MgSO$_4$ · 6 H$_2$O, MgSO$_4$ · 7 H$_2$O, MgSO$_4$ · 12 H$_2$O). The water-end portion of the diagram for the latter system is shown in Fig. 11. Careful study of this diagram will reveal that its interpretation involves merely an extension of the principles employed for interpreting the simple diagram shown in Fig. 9.

Class A: Type IV. Complete Miscibility in Solid State. Just as two liquids may dissolve in each other to form a liquid solution, so one solid may dissolve in another to form a *solid solution.* Solid, like liquid, solutions are homogeneous, and may vary in composition within wide limits. In the latter respect they differ from solid compounds whose composition is always fixed and definite. X-ray
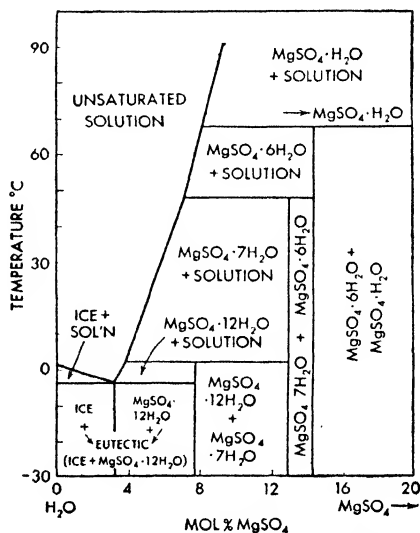
Fig. 11. Phase Diagram for the System MgSO$_4$-H$_2$O

examination of the lattices of solid solutions reveals that one constituent enters the lattice of the other and is uniformly distributed through it. This uniformity of distribution differentiates a solid solution from a mixture of solids, for in the latter instance each constituent preserves its own characteristic crystal structure.

The phase diagrams shown by the solid-liquid equilibria in binary systems where both the liquid and solid phases are completely miscible fall into three distinct groups, namely,

(1) The melting points of all solutions are intermediate between those of the pure constituents.

(2) The melting point curve exhibits a minimum.

(3) The melting point curve exhibits a maximum.

These various cases will now be discussed.

(1) *Intermediate Type Diagram.* Figure 12 shows the phase diagram for the system ammonium thiocyanate-potassium thiocyanate (NH₄CNS–KCNS) where the two con-

stituents are completely miscible in both the solid and liquid states. In this diagram the upper or *liquidus* curve gives the temperature of initial solidification, i.e., the compositions of melt saturated with various solid solutions at various temperatures, while the lower or *solidus* curve gives the temperatures at which final solidification takes place. To follow the changes involved in the solidification of any given composition of melt, consider a point such as *a* in the liquid phase. On cooling this solution, no solid phase will sepa-

Fig. 12. Phase Diagram for the System NH₄CNS–KCNS

rate until the temperature corresponding to point *b* is reached. At this point a small amount of *solid solution* of composition *c* will separate and change thereby the composition of the melt toward, say, *d*. On further cooling *d* will start freezing at *e* and the composition of the solid phase will adjust itself to *g* by solution of some ammonium thiocyanate from the melt. At all times the composition of the liquid and solid solutions at equilibrium is given by the intersections of the horizontal isothermal lines with the liquidus and solidus curves. As cooling is continued, the compositions of the melt will move along the liquidus curve toward ammonium thiocyanate while the compositions of the solid solution will move in the same direction along the solidus curve. When point *h* is reached, the solid solution in equilibrium with the small amount of melt remaining has the overall composition of the mixture, and hence at this

point the mixture solidifies completely to yield a solid solution of composition i equal to the original composition of the melt. Further cooling will result merely in lowering the temperature of i.

The same considerations apply to the cooling of any other overall composition in the diagram. Although the cooling process described above has been considered to be stepwise, actually the changes in the compositions of solid and melt involve a continuous adjustment of the concentrations to the equilibrium values called for by the liquidus and solidus curves at each temperature. It should be pointed out, however, that this mode of solidification is obtained only if the cooling rate is sufficiently slow and there is sufficiently good contact between the phases to allow equilibrium to be reached. With rapid cooling and little agitation, solid solutions of varying composition may separate, and the final congealing temperature may be lower than hi.

Since in this system no more than two phases are present at any time in equilibrium, there are no invariant points, and hence there are no horizontal portions on the cooling curves. The only discontinuities are an arrest when the liquidus curve is reached and solid starts separating, and another arrest when the overall composition line crosses the solidus curve and the melt disappears; i.e., for a composition such as a arrests are observed at temperatures corresponding to points b and i.

Continuous solid solution formation of the type discussed is fairly common. Systems besides ammonium thiocyanate-potassium thiocyanate that exhibit such behavior are lead chloride-lead bromide, silver chloride-sodium chloride, copper-nickel, cobalt-nickel, silver-gold, and naphthalene–β-naphthol.

(2) *Minimum Type Diagram.* A variation of Fig. 12 is the diagram shown in Fig. 13 for the system p-iodochlorbenzene (A)–p-dichlorbenzene (B). Here, as before, A and B form a complete series of solid solutions, but the melting point curve now, instead of being intermediate between the melting points of A and B, exhibits a minimum. Again the upper curve is the liquidus, while the lower curve is the solidus. Except for the fact that the melt composition corresponding to the minimum freezes at constant temperature to yield a solid solution, the relations in this system are similar to those described for ammonium thiocyanate-potassium thiocyanate.

Fig. 13. Phase Diagram for the System p-C_6H_4ICl–p-$C_6H_4Cl_2$

Some systems besides the one mentioned falling within this category

are sodium carbonate-potassium carbonate, potassium chloride-potassium bromide, mercuric bromide-mercuric iodide, potassium nitrate-sodium nitrate, silver-antimony, and copper-gold.

(3) *Maximum Type Diagram.* Though rarely encountered, a third possible type of diagram is one in which the melting point curve of the system exhibits a maximum. Figure 14 shows such a diagram for the system *d*-carvoxime–*l*-carvoxime ($C_{10}H_{14}N \cdot OH$). Again the liquidus is the upper curve, the solidus the lower. As in the preceding type, there is just one composition of melt which solidifies at constant temperature to yield a solid solution — the one corresponding to the maximum. All other overall compositions behave as described.

Class A: Type V. Partial Miscibility in Solid State with Eutectic. When two solids are soluble in each other to a limited degree only, *A* will dissolve a given amount of *B* to yield a

Fig. 14. Phase Diagram for the System *d*-Carvoxime–*l*-Carvoxime

saturated solution of *B* in *A*, while *B* will dissolve some *A* to yield a saturated solution of *A* in *B*. As long as these limiting concentrations are not exceeded, the solid phase is homogeneous and constitutes a single solid solution. If the range of miscibility is exceeded, however, two solid phases result, each composed of a saturated solution of one constituent in the other. It follows from the phase rule, therefore, that for equilibria between a single solid solution and melt the system will be monovariant, while for equilibria between the two solid solutions and melt the system will be invariant. Any processes involving the coexistence of the two solid solutions and melt will have to take place then isothermally.

Systems in which partial miscibility in the solid state occurs exhibit two types of diagrams. In the first of these the cooling of melt within certain composition limits results in the appearance of a eutectic involving melt and the two solid solutions. This type will be discussed in this section as type V. The other possibility is observed when the two solid solutions are not stable within certain concentration ranges, and one of these is transformed to the other through a peritectic reaction. This kind of behavior will be described as type VI.

Figure 15 shows a typical phase diagram for a system in which the partial miscibility in the solid state leads to a eutectic *B*. Line *AB* is the liquidus curve for solid solutions of silver iodide in mercuric iodide
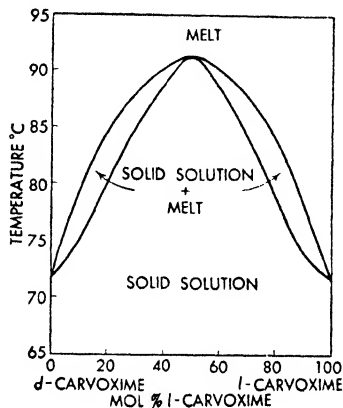
(I), while line AX is the solidus curve for the same solutions. Similarly line BC is the liquidus curve for solid solutions of mercuric iodide in silver iodide (II), while line YC is the solidus curve for these solutions. The area enclosed by the lines AX, BX, and AB contains then solid solutions I and melt, and the area enclosed by the lines BY, YC, and BC solid solutions II and melt. Below $AXYC$ no melt is present, only solid solutions. To the right of YZ and below CY the solid phase present is solid solution II, while to the left of WX and below AX the solid phase is solid solution I. The area below XY and between the lines WX and YZ is the range of partial miscibility of the solid solutions. In this area I and II coexist, with the lines WX and YZ showing the compositions of I and II respectively at each temperature. From the directions of these lines it is readily apparent that the partial miscibility of solid silver iodide and mercuric iodide decreases as the temperature is lowered. The eutectic at B involves two solid solutions with the fixed compositions given by X and Y.

Fig. 15. Phase Diagram for the System Mercuric Iodide-Silver Iodide

Among the systems exhibiting diagrams of the type shown in Fig. 15 may be mentioned silver chloride-cuprous chloride, potassium nitrate-thallium nitrate, azobenzene-azoxybenzene, naphthalene-monochloracetic acid, and the metal pairs lead-antimony, silver-copper, lead-tin, and cadmium-zinc.

Class A: Type VI. Partial Miscibility in Solid State with Peritectic. Instead of exhibiting a eutectic, two solid solutions may undergo a peritectic reaction in which a solid solution of one type is transformed to a solid solution of another type at a definite temperature. When this occurs, the phase diagram takes on the general form shown in Fig. 16 for the system iron-gold. In this diagram AD and AB are the liquidus and solidus curves, respectively, of solid solutions of gold in iron (I), while DEF and CEF are the corresponding curves for solid solutions of iron in gold (II). As may be seen from the diagram, any overall compositions lying between pure iron and point B yield only solid solutions I; those lying between pure gold and point D yield only solid solutions II. However, the solid phases resulting from mixtures between B and D depend on whether the overall composition lies between B and C, or C and D. Cooling any mixture between B and C, such as a, will yield first

solid solution I of composition f when point b is reached. Further cooling will result in further separation of I of compositions along fB, while the melt composition adjustment will proceed along bD. But when the temperature corresponding to line BD is reached, point d, solid solution I of composition B becomes unstable and begins to react with melt of composition D to form solid solution II of composition C. Since this *peritectic* reaction between the two solid solutions and melt involves the coexistence of three phases, the process must occur isothermally until the melt is all consumed and only the two solid phases remain. Once the melt has disappeared, the temperature begins to fall again, and we pass into the area in which solid solutions I and II exist, i.e., the range of partial miscibility of the two solid solutions. The lines BG and CH give the compositions of I and II respectively at various temperatures below the peritectic.
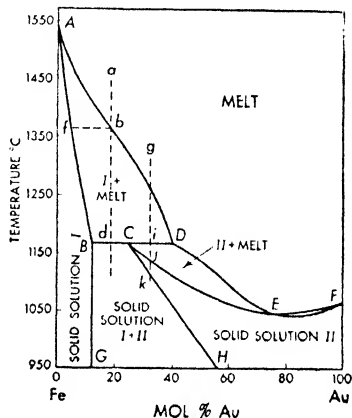
Fig. 16. Phase Diagram for the System Iron-Gold

For any composition falling between C and D, such as g, the behavior on cooling will be similar to that of a up to point i, the peritectic temperature. At i solid solution B and melt D again interact to form solid solution II of composition C; but, since now there is more melt present than is necessary to react with all of B to form C, B must disappear, and the end of the peritectic reaction must result in the presence of solid solution II and melt. This mixture of melt and II will eventually solidify at j to leave only solid solution II.

In the system under discussion the solid solution of composition j will be stable only between the temperatures corresponding to j and k. When the latter temperature is reached, j passes into the range of partial miscibility of the solid phases and breaks up into two solid solutions in equilibrium, one of gold in iron (I) and another of iron in gold (II).

The appearance of peritectics involving solid solutions is not an uncommon occurrence. Diagrams similar to Fig. 16 are found also in such binary systems as silver chloride-lithium chloride, silver nitrate-sodium nitrate, cobalt-iron, indium-thallium, and p-iodo-chlorbenzene–p-diiodobenzene.

Class B: Partial Miscibility in Liquid State. Although the discussion has centered so far on systems exhibiting a single liquid phase, there are systems in which the melt is only partially miscible over certain temperature and concentration ranges. When this separation of a

liquid into two layers occurs, the number of phases is increased, and the phase relations are thereby modified. The case to be considered here is that in which the melt is partially miscible but the solid phases are the pure constituents.

Figure 17 shows a schematic phase diagram for a system in which these conditions obtain. This figure is essentially a simple eutectic type diagram with an area of partial miscibility of the melt superimposed upon it. Outside the dome-shaped area and above the solid lines a single liquid phase is present. Within the dome-shaped area and above the line DE two liquid phases coexist whose compositions at each temperature are given by horizontal tie-lines such as bJ or kL. Since the relations below line DE have been discussed, we need elaborate only the behavior at and above this line.

For this purpose consider specifically a melt of composition a lying between D and E. Cooling of this melt will yield no new phases until at b a small amount of a second liquid layer J appears. Further cooling of these two layers to a point such as d merely changes the relative amounts of the two layers and brings about an adjustment in the concentrations along bk and JL. However, when the layer compositions reach D and E, corresponding to the overall composition point f, layer D becomes saturated with A, and the latter begins to deposit according to the scheme

$$\text{Liquid } D \longrightarrow A \text{ (solid)} + \text{Liquid } E \tag{7}$$
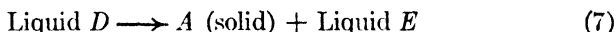
During this crystallization of A and transformation of liquid layer D into E, three phases are present in equilibrium, the system is invariant, and hence the temperature remains constant until all of D has disappeared. Once D is gone the system regains a degree of freedom, and the temperature falls to yield A in equilibrium with a single melt of composition given by EF. Final solidification takes place at F to yield a eutectic mixture of A and B.

For compositions between C and D the behavior will be somewhat different. Cooling of any such mixture will result first in a separation of A when line CD is crossed. Further cooling will shift the melt composition toward D, and when this point is reached a separation of the melt into D and E will occur. From this stage on, separation of A will proceed under isothermal conditions until D is gone, and thereafter the mixture of A and melt along EF will cool in the manner described. It

Fig. 17. Phase Diagram for System with Partially Miscible Melt

is obvious from the diagram that to the right of E only a single liquid phase is encountered, and consequently the cooling behavior is the same as that of any simple eutectic system.

In many systems the line CD is either extremely short or nonexistent. In the latter instance points D and C coincide; i.e., at the melting point of A the A-rich layer is practically pure A. Again, the eutectic point F may be displaced so close to the B axis as to obliterate practically the line FG and the $(B + \text{melt})$ area. Examples of all these various modifications of Fig. 17 may be found in the condensed binary pairs lithium-sodium, bismuth-zinc, bismuth-cobalt, chromium-copper, copper-lead, benzoic acid-water, and phenol-water.

Class C: Immiscibility in Solid and Liquid States. When two constituents are completely immiscible in each other in both the solid and liquid states, each of these substances will melt and freeze independently of the other. An example of such behavior is shown in Fig. 18 for the system vanadium-silver. Below 960° C the two elements exist as two solid phases. At 960° C the silver melts sharply to yield the liquid, which then coexists with the solid vanadium up to the melting point of the latter. At 1710° C the vanadium melts, and thereafter the system contains only the two pure liquids in two layers.
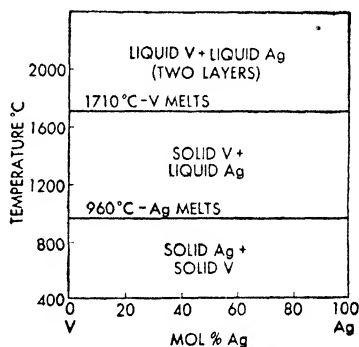
Fig. 18. Phase Diagram for the System Vanadium-Silver

Similar behavior is observed also with the metal pairs bismuth-chromium, chromium-iron, aluminum-sodium, aluminum-lead, gallium-mercury, potassium-magnesium, and others. In practically all instances the melts become partially miscible at temperatures above the fusion point of the higher melting constituent.

Composite Diagrams. As a rule binary solid-liquid diagrams are not of the simple fundamental types described, but have the elements of several types combined into a single diagram. Such a more complicated phase relation for the system magnesium-zinc is shown in Fig. 19. By applying the general principles developed for the simple types it is readily possible to interpret the more complicated equilibria occurring in this system as a combination of solid solution formation, formation of compounds, both stable ($MgZn_2$) and unstable ($MgZn$ and $MgZn_5$), and eutectic. The student should verify for himself the validity of the legends in the various areas and should sketch cooling curves for various overall compositions in the system.

Three-Component Systems. Application of the phase rule to three-component systems indicates that in such systems a single phase possesses four degrees of freedom, namely, temperature, pressure, and the compositions of two out of the three components. It is readily apparent that this number of variables presents great difficulty in the graphical presentation of the phase relations. For this reason data in ternary systems are generally presented at some fixed pressure, such as atmospheric, and at various constant temperatures. Under these conditions it is possible to present the concentration relations among the three components at any given temperature on a planar diagram. By combining such planar diagrams for various temperatures it is then possible, if desired, to construct a solid model having concentrations as a base and temperature as a vertical axis.

Fig. 19. Phase Diagram for the System Magnesium-Zinc

For a three-component system the phase rule takes the form

$$F = 3 - P + 2 = 5 - P$$

At a fixed pressure and temperature the number of degrees of freedom is reduced by two so that $F = 3 - P$, and the maximum number of phases which can occur simultaneously is thus three. This is the same number as is possible in two-component systems under a constant pressure only. It follows, therefore, that under the specified conditions an area will again indicate divariance, a line monovariance, and a point invariance.

Method of Graphical Presentation. Several schemes for plotting two-dimensional equilibrium diagrams for ternary systems have been proposed. Of these the equilateral triangle method suggested by Stokes and Roozeboom is most generally employed, and will be used here. In this method concentrations of the three components at any given temperature and pressure are plotted on an equilateral triangle such as that shown in Fig. 20. Each apex of the triangle is taken as 100 per cent of the component with which it is designated. To obtain percentages other than 100 for A, the sides AB and AC are divided into 10 (or sometimes 100) equal parts, and lines parallel to the side BC are drawn. Each of these lines represents then a definite percentage of A ranging from zero on line BC to 100 at A. Similarly lines dividing the sides BA and BC

and parallel to AC give various percentages of B, while the lines along the sides CA and CB and parallel to AB represent various percentages of C. To plot any point on the diagram such as D, having 30 per cent A, 20 per cent B, and 50 per cent C, we locate first the 30 per cent A line, namely ab, and next the 20 per cent B line, or cd. The intersection of these two lines yields the desired point. This point should lie also on the 50 per cent C line, or cf, and this fact may be utilized as a check on the accuracy of location of the point.

From the nature of the diagram it is readily apparent that the sides of the triangle represent various proportions of constituents in two-component systems. Thus lines AB, BC, and AC give the concentration relations in the binary systems $A–B$, $B–C$, and $A–C$, respectively, and
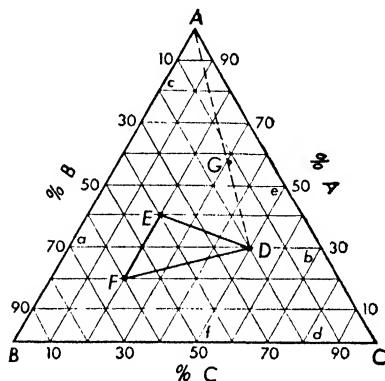
Fig. 20. Graphical Representation of Three Component Systems

any point on these lines refers only to these binary systems. On the other hand, any mixture composed of A, B, and C must lie within the diagram. In fact, the argument may be extended to show that similar relations apply also to any line or smaller triangle within the diagram. Thus all mixtures that can be prepared from D and E will lie on DE, those prepared from E and F will lie on EF, while those composed of F and D will lie on the line FD; and all possible compositions that can be prepared from D, E, and F will fall within the small triangle DEF. From the same considerations it also follows that if a point such as G lies on a straight line connecting D and A, then D must lie on an extension of the straight line through G linking A and G. This latter fact is of great importance in determining the nature of solid phases occurring in three-component systems.

Another relation to remember is this: Any mixture such as G, composed of A and D, will contain A and D in the length ratios $DG : AG$ respectively by weight, if weight percentage is plotted, or by moles, if mole percentage is plotted. By knowing the total amount of any mixture present, and by determining these lengths from the diagram, it is possible to calculate the weights (or moles) of various phases present in a given system. Such calculations find wide application in all types of separation problems involving three components.

Partially Miscible Three-Liquid Systems. Although many categories of three-component systems are possible and have been observed, prime attention will be devoted here to only two of these; namely,

(1) systems composed of three liquid components which exhibit partial miscibility to a greater or lesser degree, and (2) systems composed of two solid components and a liquid.

Systems composed of three liquids which show partial miscibility may be classified as follows:

Type I.　Formation of one pair of partially miscible liquids.
Type II.　Formation of two pairs of partially miscible liquids.
Type III.　Formation of three pairs of partially miscible liquids.

These three types will be discussed now in turn.

Type I. One Partially Miscible Pair.　Consider a pair of liquids, B and C, that are partially soluble in each other at a given temperature and pressure. If we mix relative amounts of the two so as to exceed the mutual solubility limits, two layers will be obtained, one composed of a solution of B in C, the other of C in B. Suppose we add now to the two-layer mixture a third liquid A, which is completely miscible with both B and C. Experience shows that A will distribute itself between the two layers, and will bring about thereby a greater miscibility of B and C. The increase in miscibility brought about by A depends on the amount

Fig. 21. Three Liquids with One Partially Miscible Pair

of it added and on the amount of B and C present. If sufficient A is introduced, the two layers can be changed into a single solution composed of the three liquids.

The changes in miscibility brought about by progressive additions of A to mixtures of B and C can be followed on the diagram shown in Fig. 21. Points a and b designate the compositions of the two liquid layers resulting from mixing B and C in some arbitrary overall proportion such as c, while line Ac shows the manner in which the overall composition of c is changed by addition of A. When enough A is added to change c to c_1, the compositions of the two layers are shifted from a and b to a_1 and b_1. The line a_1b_1 through c_1 connects the compositions of the two layers in equilibrium, and is called a *tie-line*. In a similar manner the compositions are changed to a_2, a_3 and b_2, b_3 when the overall compositions reach c_2 and c_3. Finally, at point b_4 sufficient A has been added to form only a single layer of this composition, and thereafter only a single solution is obtained.

It will be observed that the tie-line for composition b_4 shows that the composition of the B-rich layer at the point of complete miscibility is not

identical with b_4, but is equal to a_4. This fact indicates that complete miscibility is brought about, not by a coalescence of the two layers into one, but rather by the disappearance of the B-rich layer. Complete miscibility by the merging of the two layers into one can occur only at one point on the diagram, D. At this point the compositions of the two layers become identical, and the two solutions coalesce into a single liquid phase of constant composition. Point D is called either the *isothermal critical point* of the system, or the *plait point*, and can be obtained only by adding A to a single mixture of B and C, namely, d.

From the preceding discussion it follows that any mixture of A, B, and C of overall composition falling within the dome-shaped area will yield two liquid layers of compositions given by the appropriate tie-line through the composition of the mixture. On the other hand, any mixture of overall composition outside this area will yield only a single homogeneous solution of the three liquids. Curve aDb is frequently referred to as a *binodal* curve. In general the plait point D will fall off the maximum of the binodal curve. Furthermore, since as a rule component A will not distribute itself equally between the two layers a and b, the tie-lines will not be parallel either to BC or to each other.

Examples of systems of the type under discussion are acetic acid-chloroform-water $(A-B-C)$, and acetone-water-phenol $(A-B-C)$.

Type II. Two Partially Miscible Pairs. A system composed of three liquids such that A and B and A and C are partially miscible, while B and C are completely miscible, will exhibit a phase diagram with two
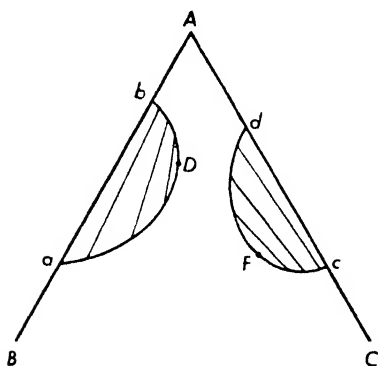
Fig. 22. Partially Miscible Liquids with Two Binodal Curves
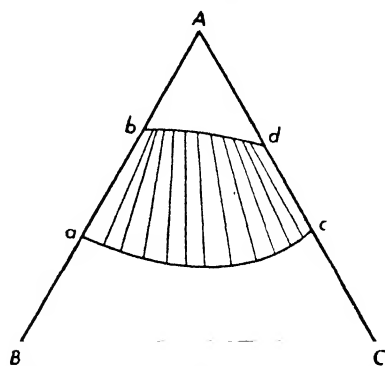
Fig. 23. Partially Miscible Liquids with Intersecting Binodal Curves

binodal curves as shown in Fig. 22. Curve aDb gives the range of compositions in which mixtures of A and B containing C are partially miscible. Again the binodal curve cFd gives the area within which C and A containing B separate into two layers. Outside the areas described by these two curves the three components are completely miscible. D and

F are the respective plait points of the two heterogeneous regions, while the indicated tie-lines show the compositions of the various layers in equilibrium. This type of diagram is shown by the liquids succinic nitrile-water-ethyl alcohol ($A–B–C$) between 18.5 and 31° C.

Although some ternary liquid systems exhibit this type of diagram at elevated temperatures, at lower temperatures, when the miscibility decreases, the two binodal curves may intersect to form the "band" type of diagram shown in Fig. 23. Here the partial miscibility area $abdc$ extends across the width of the diagram, with bd giving the composition of one layer and ac that of the other. The indicated tie-lines join solutions in equilibrium. Examples of "band" formation are found in the systems water-phenol-aniline ($A–B–C$) and water-ethyl acetate–n-butyl alcohol ($A–B–C$).

Type III. Three Partially Miscible Pairs. When all three liquids are partially soluble in each other, three binodal curves (Fig. 24) result,
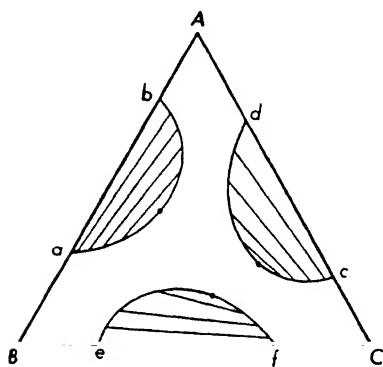
Fig. 24. Partially Miscible Liquids with Three Binodal Curves
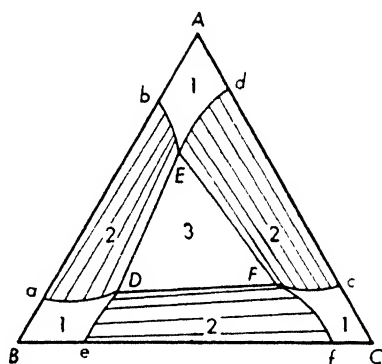
Fig. 25. Intersection of Three Binodal Curves

provided the temperature is sufficiently high to prevent intersections. Here again the dome-shaped areas indicate two-phase liquid regions, while outside these only a single phase is present. However, when two of the binodal curves intersect, as may occur at lower temperatures, the diagram takes on the appearance of Fig. 25. Here in the areas designated as 1 only a single phase exists, while in the areas marked 2 two liquid phases coexist with the equilibrium concentrations given by the connecting tie-lines. But the area marked 3 contains now *three* liquid phases. Since for three phases in equilibrium the system must be invariant at constant temperature and pressure, the compositions of the three layers must be fixed and independent of the overall composition — as long as it falls within this area. These constant compositions for the three liquid layers in equilibrium are given by the points D, E, and F.

An example of a system yielding three liquid phases in equilibrium is succinic nitrile-water-ether. At higher temperatures this system goes over to the type shown in Fig. 24.

Application of Ternary Liquid Diagrams. Ternary liquid diagrams of the type described are of considerable value in various types of separation and solvent extraction problems. In certain binary mixtures where separation of the individual pure components is desired, extraction of one component by a third solvent frequently offers advantages over distillation, crystallization, or other possible methods of segregation. This is particularly true in the removal of certain undesirable impurities from lubricating oils, where treatment with a solvent in which the impurity concentrates, but one in which the oil is either partially or totally insoluble, can bring about an improvement in the quality and properties of the lubricant. By studying the diagrams for the components in question and various solvents it is readily possible to deduce whether the separation sought can be accomplished with the given solvent, and to define the best operating conditions for optimum results.

Systems Composed of Two Solids and a Liquid. This important class of three-component systems considers the crystallization of various solid phases from solutions of two solid components in a liquid solvent. Although the solvent may be any suitable liquid in which the solid components may be soluble, attention will be focused primarily on solutions in water, which is by far our most important crystallization medium. However, the relations developed will apply equally to any other systems of the general type under discussion.

All the systems to be described below will exhibit only a single liquid phase. For one liquid phase occurring in a system, the behavior of the solid phases may be classified as follows:

> Type I. Crystallization of pure components only.
> Type II. Formation of binary compounds.
> Type III. Formation of ternary compounds.
> Type IV. Formation of complete series of solid solutions.
> Type V. Partial miscibility of solid phases.

Type I. Crystallization of Pure Components Only. When in a system composed of water and two solid components B and C only the two pure solid components appear on crystallization, the *isothermal* phase diagram has the general form shown in Fig. 26. In this diagram points D and E are solubilities in water of pure B and C respectively at the given temperature. As C is added to the solution saturated with B, the concentration of the saturated solution changes and follows the line DF. Similarly, when B is dissolved in the water solution saturated with C, the composition of the solution changes along line EF. The line DF

is, therefore, the saturation *solubility curve* of B in water containing C, while EF is the corresponding solubility curve of C in water containing B. At F, the point of intersection of the two solubility curves, the solution becomes saturated with both B and C. Since the system contains now three phases in equilibrium, B, C, and saturated solution of composition F, there are no degrees of freedom left, and hence the composition at this point must be constant. For this reason F is called the *isothermal invariant point.*

Fig. 26. Crystallization of Two Solid Components from Solution

The area above the lines DF and FE contains only unsaturated solution. However, area DFB is a two-phase region in which solid B is in equilibrium with saturated solutions of compositions lying along DF.

The particular concentration of saturated solution resulting from a given overall composition, such as G, can readily be determined from the diagram. Any mixture G of B, C, and water within the indicated area will yield at equilibrium pure solid B and a saturated solution somewhere along DF. From the relations of the equilateral triangle it follows, therefore, that the concentration of solution sought must lie on a straight line from B through G, namely, G_1. By the same token any mixture such as H will have H_1 as a saturated solution. Lines such as G_1B and H_1B which connect the concentrations of the saturated solutions with the solid phases in equilibrium with them are called tie-lines. The points of convergence of various sets of these tie-lines on a diagram determine the nature of the solid phases with which various solutions are saturated.

Corresponding to area DFB is the area EFC in which the saturating phase is C. Any mixtures falling within this area, such as I or J, yield at equilibrium solid C and saturated solutions along EF. On the other hand, the area BFC is a three-phase region. Anywhere within this area will be found solid B *and* solid C in equilibrium with saturated solution of composition F. A shift in overall composition from a point K to L or M in this area merely changes the relative proportions of B and C present from N to O or P, but the composition of F remains unaffected. Diagrams of the type under discussion are exhibited by the systems ammonium chloride-ammonium nitrate-water, sodium chloride-sodium nitrate-water, and ammonium chloride-ammonium sulfate-water.

The manner in which a diagram such as Fig. 26 may be utilized in

solving a crystallization problem may be illustrated with the following example. Suppose a mixture of the two solids B and C of overall composition Q is available and it is desired to separate from this mixture pure B. In order to recover this solid it is necessary to bring the overall composition into area DFB. This may be accomplished by the addition of water which will make the composition change along line Q–H_2O. When water is first added we enter area BFC in which both B and C are in equilibrium with solution F. However, as soon as line BF is crossed to a point such as R, all the C dissolves, and a solution R_1 is obtained saturated with solid B which can be filtered off, washed, and dried to obtain pure B. The amount of water necessary to dissolve all of C, and the yield of B, can be calculated with the aid of the diagram. From the distances QR and R–H_2O it follows that the ratio of the weight of Q to the weight of water to be added to reach R is

$$\frac{\text{Weight of } Q}{\text{Weight of water}} = \frac{R\text{–}H_2O}{RQ}$$

Knowing the original weight of Q, the weight of water to be added and the total weight of the mixture at R may be determined. Furthermore, since R is composed of B and R_1, we find in the same manner that

$$\frac{\text{Weight of } B}{\text{Weight of } R_1} = \frac{\overline{RR_1}}{\overline{BR}}$$

By using this ratio and the total weight of the mixture at R, the weight of B which can be recovered may be predicted.

An inspection of the relative proportions of B and saturated solution which are obtained by adding water to Q reveals that the maximum amount of B is recovered when R is very close to the line BF. Farther along the line the proportion of saturated solution to B is greater, and hence the yield of B is less. Another significant fact which can be learned from the diagram is that B can be recovered from mixtures of BC by the addition of pure water only when the mixture of B and C does not exceed the composition S. Once this composition is exceeded toward C, the saturation area of C is entered, and consequently only this substance may be recovered from the mixture.

The Schreinemakers " Wet Residue " Method. Experimental methods employed for the determination of phase diagrams of ternary systems containing solid and liquid phases are in general similar to those described for binary systems. Thermal analysis may be employed, but it usually yields cooling curves that are more difficult to interpret than those encountered in binary mixtures. For this reason the saturation method is used almost exclusively in the study of equilibria at and near room temperatures. The procedure for this method is briefly as follows:

Mixtures of various proportions of the solid components with water are prepared, and agitated in a thermostat until equilibrium is established. The liquid phase is separated then from the wet crystals, and both the saturated solution and the solid contaminated with liquid phase are weighed and analyzed carefully. The compositions thus obtained for both the solution and the wet residue are plotted then on a triangular diagram.

Figure 27 shows a series of points arrived at in this manner. The points S_1, S_2, etc., are the compositions of saturated solutions, while R_1,
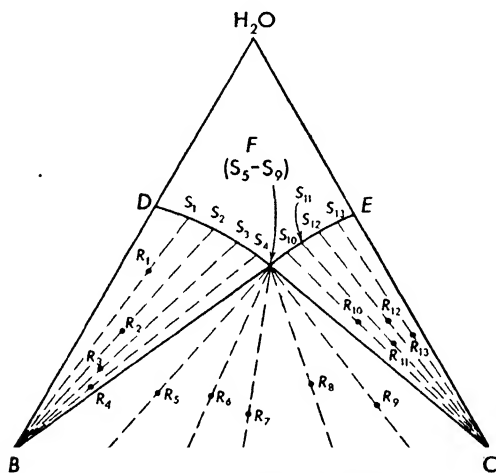
Fig. 27. Schreinemakers' Method of Wet Residues

R_2, etc., are the compositions of the corresponding wet residues. To ascertain the nature of the solid phases in equilibrium with the various solutions and present in the residues a graphical scheme is employed known as the *Schreinemakers method of wet residues.* This method is based on the fact that any wet residue composed of a given solid phase and a saturated solution must lie on a straight line *joining the composition of* the solid phase and that of the saturated solution. Consequently a tie-line drawn between any corresponding pair of R and S points must pass, on extension past R, through the composition of the solid phase. Moreover, as several solutions may have the same solid phase, all the tie-lines for such solutions must intersect at a common point, which is the composition of the common solid phase. By this reasoning we deduce that the solid phase for all solutions between D and F is B, for those between F and E is C, while F is saturated with various proportions of B and C.

Type II. Formation of Binary Compounds. In many systems composed of two salts and water compounds may be formed at a given

temperature either between the salts and water, namely, hydrates, or between the two solid salts. Although a number of possible combinations can be envisaged, only several typical cases will be discussed to indicate the nature of the diagrams obtained under these conditions.

(a) **Hydrate Formation.** Figure 28 shows the phase diagram for a system in which one of the components, B, forms a hydrate. Since the hydrate is composed of B and water, its composition must lie along the line B–H_2O and is given by D. E is the solubility of the hydrate in pure water at the given temperature, while line EF gives the solubility of the hydrate in various concentrations of water solutions containing C. Within the area DEF the hydrate D coexists with saturated solutions given by line EF. The area FGC contains again pure C in equilibrium with saturated solutions along FG. However, at F, the isothermal invariant point, the

Fig. 28. Formation of a Hydrate

solution is saturated with both D and C. As all possible mixtures of the latter two lie along line DC, the tie-lines within the area DFC must terminate along this line, as shown. Since all of the liquid phase disappears when line DC is reached, below this line only a mixture of solids D, B, and C can exist. The system sodium sulfate–sodium chloride–water at 15° C is an example of this type of behavior. The hydrate formed is $Na_2SO_4 \cdot 10\,H_2O$.

Fig. 29. Appearance of Hydrate and Pure Components

A system in which both solid components form a hydrate will have a diagram similar to Fig. 28 except that the tie-lines, instead of intersecting at C, will intercept the C–H_2O line at some point between G and C. The line DC will similarly be elevated above C. An example of such a system is magnesium chloride–calcium chloride–water at 0° C, where the hydrates $MgCl_2 \cdot 6\,H_2O$ and $CaCl_2 \cdot 6\,H_2O$ occur.

Under some temperature conditions in certain systems not only the hydrate of a component, but also the anhydrous salt, appears. The phase diagram then has the appearance of Fig. 29. Three saturation

curves are now obtained, one for each of the solid phases, and two invariant points, F and G. The solution at the invariant point F is in equilibrium with D and B, and any mixture within area DFB will yield these three phases. On the other hand, any mixture falling within area BGC will have B and C in equilibrium with the solution at the isothermal invariant point G. A diagram similar to Fig. 29 is again exhibited by the system sodium sulfate–sodium chloride–water, but at 25° C. At this temperature both Na_2SO_4 and $Na_2SO_4 \cdot 10 H_2O$ appear as saturating phases.

(b) **Double Salt Formation.** Two salts B and C, instead of forming hydrates, may react with each other to form a double salt of the general formula B_nC_m. When this happens the composition of the double salt will have to fall on the line BC, and a saturation curve for this new solid phase will have to appear in the phase triangle. The diagram for such a system is shown in Fig. 30. Here point D indicates the composition of the double salt, line FG the compositions of solutions saturated with this compound. F and G are the two isothermal invariant points. The first of these is saturated with solid B and D. The second, on the other hand, is in equilibrium with D and solid C. The significance of the remainder of the diagram should be clear from the legends in the figure.
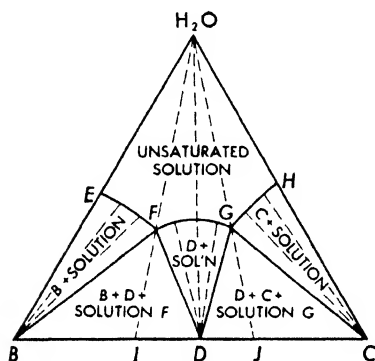
Fig. 30. Double Salt Formation

The stability of the double salt in presence of water depends on where along BC the composition of the salt falls. If point D falls between the points I and J, it is possible to prepare stable solutions of the double salt by adding water to it, as is indicated by the line D–H_2O. Such a salt is said to be *congruently saturating*. On the other hand, if the composition of the salt falls either to the left of I or to the right of J, it is impossible to prepare a saturated solution of the salt in water by adding water to D. A line from D lying between B and I to the water apex will miss curve FG to yield a mixture either within areas BDF or BEF. Similarly, for D between J and C a line to the water point will miss FG to form either a mixture of D, C, and saturated solution G, or a mixture of C and saturated solution. In either eventuality the double salt undergoes partial or complete decomposition. Double salts behaving in this manner are said to be *incongruently saturating*.

As an example of a congruently saturating double salt may be mentioned $NH_4NO_3 \cdot AgNO_3$ in the system ammonium nitrate–silver nitrate–water at 30° C. However, in the system potassium nitrate–silver nitrate–

water at the same temperature the double salt formed, $KNO_3 \cdot AgNO_3$, decomposes on addition of water, and hence is incongruently saturating.

Type III. Formation of Ternary Compounds. In certain systems not only can binary compounds be formed, but ternary compounds involving all three components as well. Figure 31 shows the phase diagram for a system in which appear a binary compound, hydrate E, and ternary compound D, composed of B, E, and water. Within area GHD compound D is in equilibrium with saturated solutions along GH, and may be recovered from these. G and H are the isothermal invariant points. G is saturated with mixtures of B and D, H with the solids D and E. All mixtures are completely solid below the lines BD and DE. Within the triangle BDE the solid phases are B, D, and E; but once line BE is crossed into area BEC, the solid phases become B, E, and C.

Fig. 31. Formation of Ternary Compound

The ternary compound D shown in Fig. 31 is evidently of the incongruently saturating type and decomposes on addition of water to it. Examples of such a compound are $CaCl_2 \cdot MgCl_2 \cdot 12\,H_2O$ (tachydrite) in the system calcium chloride-magnesium chloride-water at 25° C, and $MgSO_4 \cdot Na_2SO_4 \cdot H_2O$ in the system magnesium sulfate-sodium sulfate-water at the same temperature. On the other hand, ternary compounds such as the alums, whose general formula is $B_2SO_4 \cdot C_2(SO_4)_3 \cdot 24\,H_2O$, form congruently saturating salts which are stable in presence of water.

Type IV. Formation of Solid Solutions. When two solid components B and C are completely soluble in each other in the solid phase, a series of solid solutions ranging in composition from pure B to pure C
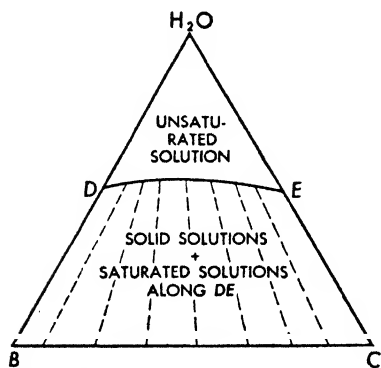
Fig. 32. Solid Solution Formation

can be recovered from a solution of these in water. Since under these conditions only two phases appear in the system, the solid solution and the liquid saturated solution, no invariant point is observed. Figure 32 shows the phase diagram for such a system. In this diagram line DE gives the compositions of saturated solutions in equilibrium with solid
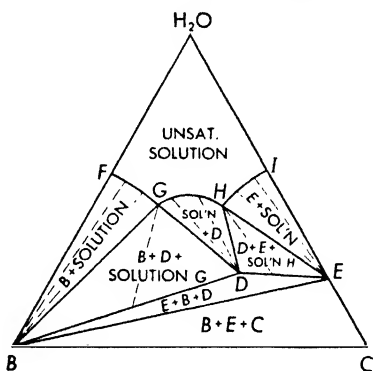
solutions of B and C of the compositions shown by the various tie-lines. In the area above DE only unsaturated solutions can be obtained. Below this line two phases occur, the saturated solutions along DE and the solid solutions in equilibrium with them.

Type V. Partial Miscibility of Solid Phases. Finally, Fig. 33 shows the phase diagram for a system in which the solid phases B and
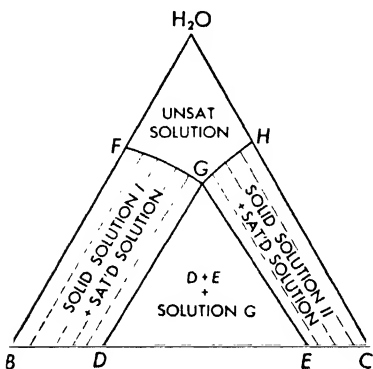
Fig. 33. Partial Miscibility of Solid Phases

C are only partially soluble in each other. Under these conditions two sets of solid solutions result, one of C in B, lying between points B and D, and another of B in C, between points E and C. Line FG gives the compositions of saturated solutions in equilibrium with the first series of these solid solutions, while line GH with the second. Between points D and E mixtures of B and C yield two solid phases of which one has the composition D, the other E. G is an isothermal invariant point such that any overall composition falling within area DGE yields this solution and the two solid solutions, D and E, in equilibrium with it.

The Three-Dimensional Phase Diagram. As was intimated before, the various triangular diagrams described heretofore are actually isothermal sections of a solid phase diagram consisting of an equilateral triangular base, along which concentrations are plotted, and a vertical axis giving temperature. Such a solid diagram can be constructed from a study of a given ternary system at various temperatures, and subsequent assembly of the isothermal sections into a three-dimensional model. The result of this kind of study for the system bismuth-tin-lead is shown schematically in Fig. 34. The relations in this system are the simplest that can be encountered, for here only the pure components appear as solid phases. In Fig. 34 points A, B, C give the melting points of pure bismuth, tin, and lead respectively. Each face of the prism indicates in turn the phase behavior of a two-component system. Thus the face ADB-Sn-Bi shows that the binary pair bismuth-tin yields a simple eutectic type diagram with the eutectic point at D. Similarly points E and F show the eutectics for the systems tin-lead and bismuth-lead. As lead is added to the binary pair bismuth-tin, the position of point D changes along line DG according to the amount of lead added. In the same manner addition of tin to F changes its locus along FG, while addition of bismuth to E changes this point along line EG. All these lines eventually intersect at G.

Lines FG, DG, and GE divide the total surface area of the prism into three distinct surfaces H, I, and J. At all temperatures above these surfaces only melt exists. As soon as a surface is reached, however, a solid phase begins to separate. On H the solid phase is bismuth, on I lead, on J tin. These surfaces represent, therefore, two-phase equilibria involving a pure component and melt. Since FG is the line of intersection of surface H, involving an equilibrium between solid bismuth and melt, and surface I, involving solid lead and melt, this line must represent three phases in equilibrium, namely, bismuth, lead, and melt. By the same reasoning along line GE, tin, lead, and melt, and along line DG bismuth, tin, and melt, are in equilibrium. When these lines eventually meet at G, four phases coexist, the three solids and melt. As for four phases in equilibrium a three-component system at constant pressure possesses zero degrees of freedom, point G must be the absolute invariant point of the

Fig. 34. The System Bi-Sn-Pb at Various Temperatures

system. This invariant point can occur only at a fixed temperature, and must correspond to a constant composition of the three components. For the system bismuth-tin-lead the coordinates of G are 97° C and approximately 51 per cent bismuth, 16 per cent tin, and 33 per cent lead. Below 97° C no melt can appear in this system, and hence below this temperature only the three solid phases coexist.

REFERENCES FOR FURTHER READING

1. W. C. Blasdale, "Equilibria in Saturated Salt Solutions," Reinhold Publishing Corporation, New York, 1927.
2. S. T. Bowden, "Phase Rule and Phase Reactions," The Macmillan Company, New York, 1938.
3. Findlay and Campbell, "The Phase Rule and Its Applications," Longmans, Green and Company, New York, 1938.
4. M. Hansen, "Der Aufbau der Zweistofflegierungen," Julius Springer, Berlin, 1936.
5. International Critical Tables, McGraw-Hill Book Company, Inc., New York, Vol. II.
6. J. S. Marsh, "Principles of Phase Diagrams," McGraw-Hill Book Company, Inc., New York, 1935.
7. A. C. D. Rivett, "The Phase Rule," Oxford University Press, Oxford, 1923.
8. H. S. Taylor, "Treatise on Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1931, Chapter IX.
9. J. E. Teeple, "The Industrial Development of Searles Lake Brines," Reinhold Publishing Corporation, New York, 1929.

PROBLEMS

1. From the following information make a rough sketch of the phase diagram for CH_3COOH, and explain what phases are in equilibrium under the conditions represented by the various areas and curves in the diagram. (a) Its melting point is $16.6°$ C under its own vapor pressure of 9.1 mm; (b) solid CH_3COOH exists in two modifications, I and II, both of which are more dense than the liquid, and I is the stable modification at low pressure; (c) phases I, II, and liquid are in equilibrium at $55.2°$ C under a pressure of 2000 atm; (d) the transition temperature from I to II decreases as the pressure is decreased; (e) the normal boiling point is $118°$ C.

2. In each of the following systems equilibrium is approached by starting with the components on the left. What is the minimum number of composition variables necessary to determine the state of each system?

   (a) $FeCl_2(s) + H_2S(g) = FeS(s) + 2 HCl(g)$

   (b) $NH_4Cl(s) = NH_3(g) + HCl(g)$

   (c) $O_2(g) + 2 NO(g) = 2 NO_2(g)$

3. In the system $NaCl-H_2O$ a simple eutectic is observed at $-21.1°$ C for a solution containing 23.3% by weight NaCl where $NaCl \cdot 2 H_2O$ and ice crystallize out of the mixture. At a composition of 27% by weight of NaCl and at $-9°$ C a peritectic point exists where the dihydrate decomposes to form anhydrous NaCl. The temperature coefficient of solubility of anhydrous NaCl is very small and positive. Make a rough sketch for the system showing clearly the phases in equilibrium in the various areas and along the various curves of the diagram.

4. The system n-heptane and 2,2,4-trimethyl pentane exhibits a simple eutectic point at $-114.4°$ C corresponding to 24 mole per cent of n-heptane [Smittenberg, Hoog, and Henkes, J. Am. Chem. Soc., **60**, 17 (1938)]. Determine analytically the maximum mole per cents of n-heptane which can be recovered by crystallization from mixtures of the two compounds containing 80, 90, and 95 mole per cent of n-heptane. *Ans.* 92.1, 96.1, and 98.0%

5. The following data are given by Andrews [J. Phys. Chem., **29**, 1041 (1925)] for the system o-dinitrobenzene–p-dinitrobenzene:

Mol % of Para Compound	Initial Melting Point,° C
100	173.5
90	167.7
80	161.2
70	154.5
60	146.1
50	136.6
40	125.2
30	111.7
20	104.0
10	110.6
0	116.9

Construct a temperature-composition diagram for the system, and determine therefrom the eutectic temperature and composition.

6. Using the plot constructed in the preceding problem, find graphically the maximum percentage of p-dinitrobenzene which can be recovered pure by crystallization from mixtures of the two compounds containing originally 95%, 75%, and 45% of the para compound.

7. From the phase diagram of the system Cd-Bi given in the text, estimate the solubility of Cd in Bi at $200°$ C. *Ans.* 110.5 g/100 g Bi

8. Interpret the phase diagram, Fig. 35, for the system Al-Ni, explaining what phases are in equilibrium under the various conditions represented by the areas and curves.

9. In the preceding problem sketch cooling curves showing the complete solidification of melts having compositions a, b, and c.
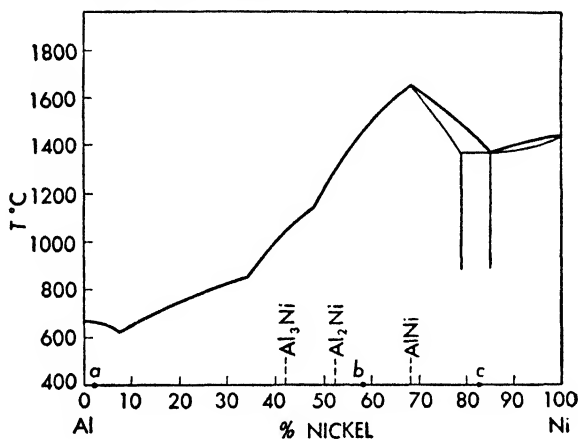
Fig. 35

10. Will the addition of a small amount of impurity to a pure compound or element always lower its melting point? Explain your answer.

11. In Fig. 7 sketch and interpret the cooling curves which would result when melts corresponding to 20, 40, and 60 mol per cent of $FeCl_3$ are cooled.

12. By referring to Fig. 9, state the conditions of temperature and composition which must be met in order to crystallize in the pure state the compound $CaF_2 \cdot CaCl_2$ from the CaF_2-$CaCl_2$ system.

13. By referring to Fig. 11, explain how you would proceed to obtain the optimum amount of pure $MgSO_4 \cdot 6 H_2O$ crystals from a dilute aqueous solution of $MgSO_4$.

14. The system SO_3-H_2O exhibits congruent melting points at the compositions by weight of 68.96%, 81.63%, and 89.89% SO_3. What are the formulas of the corresponding compounds?

15. CCl_4 and H_2O may be assumed to be immiscible. Apply the phase rule to the system H_2O-CCl_4-I_2, and explain what variables must be specified in order to determine the state of the system.

16. The following data are given by Prutton, Brosheer, and Maron [J. Am. Chem. Soc. **57**, 1656 (1935)] for the system NH_4Cl-NH_4NO_3-H_2O at 25° C:

Saturated Solution		Wet Residue	
NH_4NO_3	NH_4Cl	NH_4NO_3	NH_4Cl
%	%	%	%
67.73	0	—	—
66.27	2.00	88.20	0.79
64.73	3.82	88.00	1.34
62.24	5.58	90.25	1.65
61.68	6.97	87.65	2.28
53.49	11.08	23.31	62.22
36.99	15.80	13.63	66.29
19.05	21.81	7.09	72.75
0	28.33	—	—

Construct and interpret the phase diagram for the system, and determine the ternary composition. Is there evidence of hydrate or double salt formation in this system?

17. Using the plot constructed in the preceding problem, determine the maximum theoretical recovery at 25° C of NH_4Cl from a dry salt mixture of NH_4Cl-NH_4NO_3 containing 80% by weight of NH_4Cl.   *Ans.* 96.1%

18. For the system KNO_3-$NaNO_3$-H_2O a ternary point exists at 5° C at which the two anhydrous salts are in equilibrium with a saturated solution containing 9.04% by weight of KNO_3 and 41.01% $NaNO_3$. Determine analytically the maximum weight of KNO_3 which could be recovered pure from a salt mixture containing 70 g of KNO_3 and 30 g of $NaNO_3$ by crystallization from an aqueous solution at 5° C.   *Ans.* 87.5%
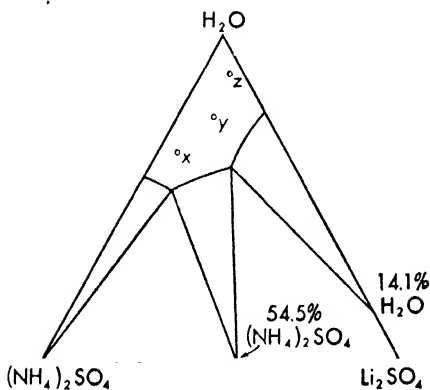
Fig. 36

19. Interpret the phase diagram, Fig. 36, for the system $(NH_4)_2SO_4$-Li_2SO_4-H_2O at 25° C.

20. In the diagram of the preceding problem, what will be the composition of the first crystals formed by the evaporation of H_2O from solutions of compositions x, y, and z at 25° C?

21. The following hydrates may under proper conditions be crystallized from the system $Ba(OH)_2$-$NaOH$-H_2O at 30° C: $Ba(OH)_2 \cdot 8 H_2O$, $Ba(OH)_2 \cdot 3 H_2O$, $Ba(OH)_2 \cdot H_2O$, and $NaOH \cdot H_2O$. Sketch and interpret a phase diagram for this system.

22. The following data are for the system water-alcohol-benzene at 25° C. The first two columns give the percentages by weight of alcohol and benzene in one layer, while the third column gives the percentage by weight of water in the layer conjugate to these.

Layer I		Layer II
% C_6H_6	% C_2H_5OH	% H_2O
1.3	38.7	
9.2	50.8	
20.0	52.3	3.2
30.0	49.5	5.0
40.0	44.8	6.5
60.0	33.9	13.5
80.0	17.7	34.0
95.0	4.8	65.5

Construct a phase diagram for the system, and draw in the tie-lines.

23. An aqueous solution contains 46% by weight of ethyl alcohol. Using the diagram of the preceding problem, find how much alcohol would be extracted from 25 g of this solution by 100 g of C_6H_6.   *Ans.* 5.05 g

Electrolytic Transference and Conductance

Introduction. The branch of physical chemistry that concerns itself with the interrelation of chemical phenomena and electricity is called *electrochemistry.* Since according to our present concepts all matter is electrical in nature, electrochemistry could well encompass the entire field of chemical reactions. However, in a more modest way this subject confines itself primarily to a study of the electrical properties of solutions of electrolytes, and to the elucidation of the relation between chemical action and electricity in such systems. Even within these confines the phenomena encountered are of such theoretical and practical importance that this and the three chapters following are to be devoted to an exposition of various aspects of the subject.

Ohm's Law and Electrical Units. The strength of an electric current flowing through a conductor is determined by the difference in potential applied across the conductor and by the resistance offered by the conductor to the current. According to Ohm's law the relation among these three quantities is given by

$$I = \frac{\mathcal{E}}{R} \tag{1}$$

where I is the strength of a current flowing through a resistance R under an applied potential $\mathcal{E}$. From this equation it is evident that the current strength is directly proportional to the difference in potential and inversely proportional to the resistance. By an appropriate choice of units the constant of proportionality is made unity.

The quantities I, R, and $\mathcal{E}$ may be expressed in *electromagnetic* units, based on the law of attraction or repulsion of magnets, in *electrostatic* units, based on the Coulomb law of force between electric charges, or in *absolute practical* units, deduced from the above. However, for convenience of comparison and standardization a system of units, known as the *international*, has been defined, and is the one generally employed in electrochemistry. In this system the unit of current, or rate of flow of electricity, is the *ampere*, which is defined as an invariable current of such strength that on passage through a water solution of silver nitrate

it will deposit 0.00111800 gram of silver in 1 sec. In turn, the unit of resistance, the *ohm*, is defined as the resistance at 0° C of a column of mercury of uniform cross section, 106.300 cm long, and containing 14.4521 grams of mercury. From these two units the unit of potential, the *volt*, follows from Ohm's law as the potential difference required to send a current of 1 amp through a resistance of 1 ohm.

The unit of quantity of electricity, or charge, is the *coulomb*. A coulomb is the quantity of electricity necessary to deposit 0.00111800 gram of silver from a solution of silver nitrate, no matter what the time is. Consequently, a current of 1 amp is one which transports 1 coulomb of electricity per second. Since the quantity of electricity carried by a current must equal rate of transport times the time, the charge Q carried by a current I in t sec must be

$$Q = It \tag{2}$$

coulombs. Another unit of quantity of electricity which we shall employ frequently is the *faraday*, $\mathcal{F}$. A faraday is equal to 96,500 coulombs.

The electrical work w performed when a current of strength I passes for t sec through a resistance across which the potential drop is $\mathcal{E}$ is given by Joule's law, namely,

$$w = \mathcal{E}It = \mathcal{E}Q \tag{3}$$

where w is expressed in *joules*. The *joule* is the electrical unit of energy, and is defined as the amount of work performed by a current of 1 amp flowing for 1 sec under a potential drop of 1 volt. The work in joules is readily convertible to other energy units through the relations

$$1 \text{ joule} = 1 \times 10^7 \text{ ergs} \tag{4}$$
$$= 0.2389 \text{ cal}$$

Finally, the rate at which work is being done by an electric current is expressed in *watts*. A *watt* is work performed at the rate of 1 joule per second, and is obviously a unit of electrical power. From equation (3) the power in watts p delivered by a current follows as

$$p = \mathcal{E}I = \frac{\mathcal{E}Q}{t} \tag{5}$$

A larger unit of power is the kilowatt, which is equal to 1000 watts.

Electronic and Electrolytic Conductors. All substances may in general be subdivided into conductors or nonconductors of electricity. The latter are frequently referred to also as *dielectrics* or *insulators*. The distinction between these two classes is not so sharp as the classification would seem to indicate. There is rather a shading off in conductivity as we pass from one class to the other, making it sometimes very difficult

to say whether a substance is to be considered a poor conductor or an insulator. This may be seen from Table 1, where the specific resistivity ρ for a number of substances is given. The substances with low resistivity are good conductors, while those with high resistivity are poor conductors or dielectrics.

TABLE 1

SPECIFIC RESISTIVITY OF VARIOUS SUBSTANCES

Substance	Temp. (°C)	ρ (Ohm-cm)	Substance	Temp. (°C)	ρ (Ohm-cm)
Silver	20	1.59×10^{-6}	Water (pure)	20	1×10^{6}
Copper	20	1.72×10^{-6}	Slate	—	1×10^{8}
Magnesium	20	4.6×10^{-6}	Celluloid	16	4×10^{10}
Mercury	20	9.58×10^{-5}	Wood	—	$10^{10} - 10^{13}$
Bismuth	20	1.20×10^{-4}	Glass (ordinary)	20	9×10^{13}
Graphite	0	8.00×10^{-4}	Sulfur	17	8×10^{15}
CaCl$_2$ (fused)	750	0.86	Paraffin oils	—	1×10^{16}
H$_2$SO$_4$ (1 M)	20	1.31			
KCl (0.01 M)	20	7.83×10^{2}			
Acetic acid (0.001 M)	18	2.44×10^{4}			

Flow of electricity through a conductor involves a transfer of electrons from a point of higher negative potential to one of lower. However, the mechanism by which this transfer is accomplished is not the same for all conductors but depends on their nature. On the basis of mechanism of current flow conductors may be classified into two types, namely, (a) electronic, and (b) electrolytic conductors. In electronic conductors, of which solid and molten metals and certain solid salts (cupric sulfide, cadmium sulfide) are examples, conduction takes place by direct migration of electrons through the conductor under the influence of an applied potential. Here the atoms or ions composing the conductor are not involved in the process, and, except for a vibration about their mean position of equilibrium, they remain stationary. The only noticeable effect produced by the current in the conductor is a heating resulting from the dissipation of electrical energy in the form of heat. Another characteristic of electronic conductors is their increase in resistance as the temperature is raised. Thus, the specific resistance of platinum increases from 1×10^{-7} at $-265°$ C to 1×10^{-5} at $20°$ C, and to 2.54×10^{-5} at $400°$ C. This fact is utilized in platinum resistance thermometers for estimating temperature by observing the change in resistance with change in temperature. The decrease in resistance becomes especially pronounced at temperatures of 3 or 4° K. Under these conditions the resistance of metals becomes so low that there is practically no opposition left to electron flow. In this state the metals are said to be *superconducting*.

Electrolytic conduction is encountered in solutions of strong and weak electrolytes, in fused salts, and also in some solid salts, such as sodium chloride and silver nitrate. In distinction to electronic conductors, electron transfer in electrolytic conductors takes place not by the flow of free electrons, but by a *migration of ions*, both positive and negative, toward the electrodes. This migration involves not only a transfer of electricity from one electrode to the other, but also a transport of matter from one part of the conductor to another. Further, current flow in electrolytic conductors is always accompanied by chemical changes at the electrodes which are quite characteristic and specific for the substances composing the conductor and the electrodes. Finally, while the resistance of electronic conductors increases with temperature, that of electrolytic conductors always decreases as the temperature is raised.

Fig. 1. Electrolytic Conduction

Mechanism of Electrolytic Conduction. The mechanism by which an electric current passes through a solution, with the attendant ionic migrations and chemical reactions, can best be understood from a specific example. For this purpose consider a cell, Fig. 1, composed of two inert electrodes, in this case platinum, connected to a source of current B and dipping into an aqueous solution of sodium chloride. The electrode C, connected to the negative side of B, is called the *cathode*. This is the electrode by which electrons from B, say a battery, enter the solution. In turn, electrode A, connected to the positive side of the battery, is termed the *anode*. It is the electrode through which the electrons leave the solution to return to B. In solution we have sodium ions, chloride ions, and also some hydrogen and hydroxyl ions due to the very slight ionization of the water. Now, when the circuit is closed and a current passes through the solution, it is found that chlorine gas escapes at the anode and hydrogen gas at the cathode, while sodium hydroxide forms in the solution immediately adjacent to the cathode. These changes may be explained as follows. Electrons enter the solution at the cathode C by combining with hydrogen ions in solution to form monatomic hydrogen. Two atoms of hydrogen thus deposited on the electrode combine then to form a molecule of hydrogen which escapes from the electrode as a gas. The reactions involved can be represented by
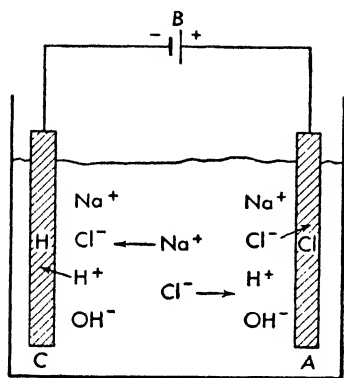
$$2\,H^+ + 2\,\ominus\,\text{(electrons)} = 2\,H$$
$$2\,H = H_2(g)$$
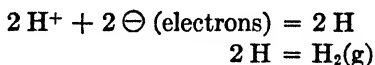
Again, electrons leave the solution at the anode by the discharge of chloride ions, with each chloride ion giving up one electron to the electrode and becoming a chlorine atom. The electrons thus liberated flow through the external circuit from the electrode to the source of potential, while the nascent chlorine atoms combine with each other to form gaseous chlorine which escapes. The reactions here are

$$2 \ Cl^- = 2 \ Cl + 2 \ominus$$
$$2 \ Cl = Cl_2(g)$$
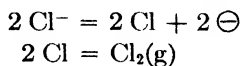
We see, therefore, that in terms of this picture two electrons are removed from the cathode to form a molecule of hydrogen, and simultaneously two electrons are given up to the anode by chloride ions to form a molecule of chlorine. The net result is a transfer of two electrons from the cathode side of the circuit to the anode side.

How the electrons get from the cathode to the anode may be surmised from the fact that negative ions, or *anions*, migrate when the circuit is closed toward the anode, while positive ions, or *cations*, migrate toward the cathode. As these particles are charged, their motion constitutes a flow of electricity, or an electric current. The anions move to the anode, and hence electrons are carried by these ions from the cathode to the anode. Again, since transport of positive electricity to the *left* may be considered a flow of negative electricity to the *right*, the migration of cations to the cathode is equivalent to a flow of electrons in the opposite direction. Consequently, the net result of the migration is a flow of electrons through the solution in the direction of the current, with each ion carrying part of the current and thus contributing its share to the transport of electricity through the solution.

The formation of sodium hydroxide in the cathode portion of the cell is understood when we remember that from this part of the solution hydrogen ions have been removed by discharge on the electrode, leaving in solution an excess of hydroxyl ions. Since these have no existence independent of positive ions, sodium ions migrate into the cathode compartment in quantity just sufficient to give an electrically neutral solution. When this solution of sodium and hydroxyl ions is evaporated, solid sodium hydroxide is obtained.

The process of current passage through an electrolytic conductor with all the accompanying chemical and migratory changes is called *electrolysis*. From the above discussion the mechanism of electrolysis may be summarized by saying that, (1) electrons enter and leave the solution through chemical changes occurring at the electrodes, and (2) electrons pass through the solution by migration of ions. Just as many electrons pass through the solution and leave it as entered it, no more, no less. The evidence for this last statement will follow from Faraday's laws of electrolysis.

Faraday's Laws of Electrolysis. The chemical reaction which occurs during electrolysis at the anode need not be necessarily a deposition of ions, but may be any *oxidation* reaction which may take place at the electrode, as solution of a metal or oxidation of ferrous to ferric ions. Similarly, the reaction at the cathode may be any *reduction* possible under the circumstances, as solution of iodine to form iodide ions, or reduction of stannic to stannous tin. Still, no matter what the nature of the reaction may be, Michael Faraday found that the *mass of a substance involved in reaction at the electrodes is directly proportional to the quantity of electricity passed through the solution.* This statement is known as *Faraday's first law of electrolysis.* The law has been shown to hold very rigidly provided the passage of electricity takes place entirely by electrolytic conduction. It applies to molten electrolytes as well as to solutions of electrolytes, and is independent of temperature, pressure, or the nature of the solvent, as long as the latter can promote ionization of the solute.

From Faraday's first law of electrolysis the quantity of electricity necessary to deposit one equivalent weight of silver may readily be calculated. Since by definition 1 coulomb deposits 0.00111800 gram of silver, and since the mass deposited is directly proportional to the quantity of electricity, the number of coulombs required for deposition of 1 gram atomic weight of silver, 107.88 grams, must be

$$\frac{107.88}{0.00111800} = 96,494 \text{ coulombs}$$

The question that immediately arises is: What mass of other substances will this quantity of electricity deposit or form? If two cells, one composed of silver electrodes in silver nitrate, the other of copper electrodes in copper sulfate, are connected in series, any current passed through one cell must also pass through the other. With such a setup it can be shown that the weight of copper deposited per 0.001118 gram of silver, i.e., per coulomb of electricity, is 0.0003294 gram. From this figure the quantity of electricity necessary to deposit 1 gram atomic weight of copper, 63.57 grams, follows as

$$\frac{63.57}{0.0003294} = 192,988 = 2(96,494) \text{ coulombs}$$

However, since copper is divalent, an equivalent is 63.57/2 grams. Consequently, to deposit *one equivalent* of copper only one-half the above amount of electricity is required, or again 96,494 coulombs.

From a series of such experiments Faraday arrived at his *second law of electrolysis*, namely, that *the masses of different substances produced during electrolysis are directly proportional to their equivalent weights.*

Another way of stating this law is that *the same quantity of electricity will produce chemically equivalent quantities of all substances* resulting from the process. Moreover, since we saw above that 96,494 coulombs will yield one equivalent of silver, a direct consequence of Faraday's second law is that *during electrolysis 96,494 coulombs of electricity will yield one equivalent weight of any substance.* To honor the man who discovered these laws, the name faraday (algebraic symbol $\mathcal{F}$) has been given to this quantity of electricity, i.e., 1 faraday = 96,494 coulombs. For ordinary calculations 1 faraday = 96,500 coulombs will be sufficiently exact.

It is essential to realize that 1 faraday will produce one equivalent of *each* of the primary products of electrolysis. By a primary product is understood one that is formed directly by the current rather than by subsequent chemical action. To the latter Faraday's laws do not apply. For instance, 1 faraday passed through a sodium chloride cell will produce one equivalent each of chlorine, hydrogen, and sodium hydroxide. These are the primary products of the electrolysis. However should the chlorine happen to diffuse into the cathode compartment and react with the sodium hydroxide to form sodium hypochlorite and sodium chlorate, the amounts of these formed will not depend on the quantity of electricity passed through the cell, but rather on the operating conditions. In fact, with proper care the formation of these *secondary* products can be avoided, but not that of the primary.

By applying Faraday's laws, the weight of primary products formed in any electrolytic process may be calculated very simply from a knowledge of the strength of the current and its time of passage. Thus, suppose a solution of silver nitrate is electrolyzed between silver electrodes with a current of 0.20 amp flowing for 30 min. The quantity of electricity delivered to the cell is evidently $0.2 \times 30 \times 60 = 360$ coulombs, and hence the weight of silver deposited is 360/96,500 of an equivalent, or

$$\frac{107.88 \times 360}{96,500} = 0.4025 \text{ g}$$

Significance of Faraday's Laws. Like the first law of electrolysis, the second law holds very rigidly for electrolytes both in solution and when fused. Its validity is again independent of temperature, pressure, and the nature of the solvent. A reason for the exactness of these laws, and an insight into their significance, may be found from the following simple calculation. Millikan and others have established that the charge on the electron is equal to 1.60×10^{-19} coulombs of electricity. Consequently the number of electrons constituting 1 faraday is

$$\frac{96,494}{1.60 \times 10^{-19}} = 6.03 \times 10^{23} \text{ electrons}$$

But, 6.03×10^{23} is exactly Avogadro's number. Hence, we must conclude that 1 faraday of electricity is associated with 6.03×10^{23} particles of unit charge, or, in general, with one equivalent of a chemical substance. One equivalent of positive ions lacks this number of electrons, while one equivalent of negative ions has this number of electrons in excess. When 1 faraday of electricity is passed through a solution, Avogadro's number of electrons is removed from the cathode by reduction of one equivalent of substance, and exactly the same number of electrons is donated to the anode as a result of some oxidation. Through the solution the equivalent of Avogadro's number of electrons is carried by the migration of positive and negative ions to the appropriate electrodes. In this way the whole process of electrolysis reduces itself to the transport of a given number of electrons through the electrolytic conductor. Since a given number of electrons does not depend on pressure, temperature, or other such factors, neither should the process of electrolysis, and this is the case.

Coulometers. As Faraday's laws are obeyed so rigidly, they may be utilized to determine the quantity of electricity passing through a circuit by observing the chemical changes produced by the same current in a suitable electrolytic cell. A cell used for this purpose is called a *coulometer.* The coulometer is placed in the circuit in series with any other apparatus, and is allowed to remain there as long as the current is flowing. It is then removed, and the chemical changes produced by the current are estimated by some appropriate means.

The coulometer commonly employed for precise work is the silver coulometer. This coulometer consists of a platinum dish serving as both cathode and cell vessel, and pure silver as anode. The electrolyte is an aqueous solution of purified silver nitrate. The dish is weighed before electrolysis, and the cell is assembled. After electrolysis the electrolyte is carefully decanted, the deposit of silver on the dish is thoroughly washed with distilled water, and then dish and silver are dried and weighed. From the increase in weight due to deposited silver the quantity of electricity passed through the coulometer can readily be calculated. As an added precaution the anode is separated from the dish by a porous cup to prevent any silver which may chip off the electrode from falling into the dish. With care such coulometers can give results accurate to 0.05 per cent or better.

Another coulometer which has been shown to be capable of very high precision is the iodine coulometer.[1] In this coulometer the iodine liberated from a solution of potassium iodide by passage of current is estimated by titration with sodium thiosulfate or arsenious acid. For less

[1] For details on this and the other coulometers, see the book by MacInnes listed at end of the chapter.

accurate work copper coulometers, consisting of copper electrodes in a solution of coper sulfates, are quite suitable. Here the copper deposited is estimated by weighing.

Transference and Transference Numbers. Although current is transported through a solution by migration of positive and negative ions, the fraction of the total current carried by each is not necessarily the same for the two ions. Thus experiments show that in dilute solutions of magnesium sulfate the magnesium ion carries about 0.38 of the total current, while the sulfate ion carries the balance, or 0.62. Similarly in dilute nitric acid solutions the nitrate ion carries only 0.16 of the total current, the hydrogen ion 0.84. The sulfate and hydrogen ions transport the greater fraction of the total current because in their respective solutions they move faster than the other ions present. If both ions in a solution moved with the same speed, each would transfer past any fixed plane in the solution the same quantity of electricity in any given time. However, when the speeds of two ions are not the same, in any given period of time the faster ion will carry past any plane a greater fraction of the current, and hence will perform a greater percentage of the total work involved in current transfer.

The quantitative relation between the fraction of the current carried by an ion and its speed can be established as follows. Consider two parallel plates, 1 cm apart, Fig. 2, across which is applied a potential of 1 volt, and between which is contained some volume of an electrolyte. Let the velocity of the cation in this solution be v_+ cm per second, the charge of the ion be z_+, and the number of these ions n_+. Similarly, let the velocity of the anion be v_-, its charge z_-, and the number of these n_-. Then the quantity of electricity transported by the cation in 1 sec,

Fig. 2. Relation of Current to Ionic Velocities

i.e., the current due to the cation, is going to be all the electricity possessed by all the cations that lie within a distance v_+ of the negative plate, or, in other words, all those which lie within the volume $ABCDE-FGH$. The number of these ions is obviously the fraction v_+ of the total or n_+v_+. Again since the charge of each ion is z_+, and the quantity of electricity associated with unit charge is e, the electronic charge, the current carried by the positive ions, must be

$$I_+ = n_+v_+z_+e \tag{6}$$

In a like manner, the current carried by the anions to the positive plate

must be that corresponding to all the electricity possessed by the anions in the volume $A'B'C'D'E'F'G'H'$, namely,

$$I_- = n_- v_- z_- e \tag{7}$$

Consequently, the total current carried by both ions is

$$I = I_+ + I_- = n_+ v_+ z_+ e + n_- v_- z_- e \tag{8}$$

But, the condition for electroneutrality of the solution demands that the total charge of the cations must be equal to that of the anions, namely,

$$n_+ z_+ = n_- z_- \tag{9}$$

Therefore,
$$\begin{aligned} I &= n_+ z_+ e v_+ + n_+ z_+ e v_- \\ &= n_+ z_+ e (v_+ + v_-) \end{aligned} \tag{10}$$

From equations (6) and (10) the fraction of the total current carried by the cations, t_+, follows as

$$\begin{aligned} t_+ &= \frac{I_+}{I} = \frac{n_+ v_+ z_+ e}{n_+ z_+ e (v_+ + v_-)} \\ &= \left(\frac{v_+}{v_+ + v_-} \right) \end{aligned} \tag{11}$$

while the fraction of the total current carried by the anions, t_-, follows from equations (7) and (10) as

$$\begin{aligned} t_- &= \frac{I_-}{I} = \frac{n_- v_- z_- e}{n_+ z_+ e (v_+ + v_-)} = \frac{n_+ z_+ v_- e}{n_+ z_+ e (v_+ + v_-)} \\ &= \left(\frac{v_-}{v_+ + v_-} \right) \end{aligned} \tag{12}$$

t_+ and t_- are called the *transport* or *transference numbers* of the cation and anion respectively. These numbers give the fraction of the total current carried by a given ion in a solution. On dividing equation (11) by equation (12) we see that

$$\frac{t_+}{t_-} = \frac{v_+}{v_-} \tag{13}$$

and hence the transport numbers of the ions, and therefore the fractions of the total current they carry, are directly proportional to their absolute velocities. When these are equal, $t_+ = t_-$ and both ions contribute equally to the transport of current. However, when v_+ does not equal v_-, t_+ will not equal t_-, and the two ions will carry different proportions of the total current. Still, no matter what the ratio between t_+ and t_- may be, since the two ions carry between them all of the current, the sum of the two transference numbers must be unity, i.e.,

$$t_+ + t_- = 1 \tag{14}$$

through the cell, it is possible to calculate the transport number of the hydrogen ion as follows:

Suppose that the volume swept out by the boundary in moving from aa' to bb' during the passage of Q faradays of electricity, as measured in the coulometer, is V cc. Then, if the concentration of the acid is C equivalents per liter, the number of equivalents of acid carried toward the cathode by the current is $V \times C/1000$ equivalents. Since this number of equivalents is carried toward the cathode by hydrogen ions, and as the total current carried is Q equivalents, obviously the transport number of the hydrogen ion, t_+, is given by

$$t_+ = \frac{V \times C}{1000\, Q} \tag{18}$$

Although as described here the method appears to be very simple, actually considerable care and practice are necessary to obtain precise results. Further, certain conditions must be met in choosing the solutions to be employed in connection with the one to be studied, and some corrections must be applied for the change in volume due to deposition or solution at the electrodes. For details of these various factors the student must be referred to the excellent discussion given by MacInnes,[1] who with Longsworth has made very valuable contributions to the technique and precision of this method. These workers have shown that, with strict observance of the factors alluded to and with the appropriate equipment, results can be obtained which not only agree with the Hittorf method, but may actually exceed it in accuracy.

Results of Transference Measurements. In Table 2 are given the transference numbers of the cations for a number of electrolytes in aqueous solution. As far as can be observed, these numbers are not affected by current strength. They vary somewhat with concentration, but the variation is, in general, not large. As a rule the transference numbers that are large in dilute solution increase with rise in concentration, while those which are small decrease. However, there are exceptions to this rule. With increase in temperature the transport numbers of the cation and anion tend to equalize and approach a value of 0.5. Hence transference numbers greater than 0.5 decrease as the temperature is raised, while those less than 0.5 increase. This variation is illustrated in the data for hydrochloric acid and potassium chloride given in the table.

Electrolytic Conductance. The resistance of an electrolytic conductor to current passage can be determined by the application of Ohm's law to such conductors. However, instead of the resistance, it is customary to speak of the *conductance*, which is merely the *reciprocal* of the electrical resistance.

[1] See reference at end of chapter.

TABLE 2

TRANSFERENCE NUMBERS OF CATIONS [1]

Substance	Temp. (° C)	Concentration — Equivalents per Liter							
		0.005	0.01	0.02	0.05	0.10	0.20	0.50	1.00
HCl	0	0.847	0.846	0.844	0.839	0.834	—	—	—
	18	0.832	0.833	0.833	0.834	0.835	0.837	0.840	0.844
	25	—	0.825	0.827	0.829	0.831	0.834	—	—
	96	—	0.748	—	—	—	—	—	—
HNO₃	20	0.839	0.840	0.841	0.844	—	—	—	—
H₂SO₄	20	—	—	0.822	0.822	0.822	0.820	0.816	0.812
NH₄Cl	25	—	0.491	0.491	0.491	0.491	0.491	—	—
AgNO₃	25	—	0.465	0.465	0.466	0.468	—	—	—
LiCl	25	—	0.329	0.326	0.321	0.317	0.311	—	—
KCl	25	0.490	0.490	0.490	0.490	0.490	0.489	—	—
	30	0.498	0.498	0.498	0.498	0.497	0.496	—	—
NaCl	25	—	0.392	0.390	0.388	0.385	0.382	—	—
NaOH	25	—	0.203	—	0.189	0.183	0.177	0.169	0.163
NaC₂H₃O₂	25	—	0.554	0.555	0.557	0.559	0.561	—	—
CaCl₂	25	—	0.426	0.422	0.414	0.406	0.395	—	—
CdSO₄	18	—	0.389	0.384	0.374	0.364	0.350	0.323	0.294
CuSO₄	18	—	—	0.375	0.375	0.373	0.361	0.327	—

As is well known, the resistance of any conductor is proportional directly to its length and inversely to its cross-sectional area, namely,

$$R = \rho \frac{l}{A} \qquad \qquad (19)$$

where R is the resistance in ohms, l the length in centimeters, and A the area in square centimeters. The proportionality constant ρ, termed the *specific resistivity*, is the resistance of a conductor 1 cm in length and with a cross-sectional area of 1 sq cm. The value of ρ depends on and is characteristic of the nature of the conductor. From equation (19) the expression for the corresponding conductance L follows as

$$L = \frac{1}{R} = \frac{1}{\rho}\left(\frac{A}{l}\right)$$
$$= L_s\left(\frac{A}{l}\right) \qquad (20)$$

where $L_s = 1/\rho$ is the *specific conductance* of the conductor. This quantity may be considered the conductance of 1 cm cube (not cc) of material, and is expressed in reciprocal ohms or *mhos*.

[1] International Critical Tables, vol. VI, p. 309; Longsworth, J. Am. Chem. Soc., **54**, 2741 (1932); **57**, 1185 (1935).

Although the specific conductance, like the specific resistivity from which it is derived, is a property of the conducting medium, in dealing with solutions of electrolytes a quantity of greater significance is the *equivalent conductance* Λ. The equivalent conductance of an electrolyte is defined as the conductance of a volume of solution containing one equivalent weight of dissolved substance when placed between two parallel electrodes 1 cm apart, and large enough to contain between them all of the solution. Λ is never determined directly, but is calculated from the specific conductance. If C is the concentration of a solution in gram equivalents per liter, then the concentration per cubic centimeter is $C/1000$, and the volume containing one equivalent of the solute V_e in cubic centimeters is therefore,

$$V_e = \frac{1}{\left(\dfrac{C}{1000}\right)} = \frac{1000}{C}$$

Since L_s is the conductance of a centimeter cube of the solution, the conductance of $1000/C$ cc, and hence Λ, will be

$$\Lambda = L_s V_e = \frac{L_s 1000}{C} \tag{21}$$

Equation (21) is the defining expression for the equivalent conductance. It must be remembered that C in this equation is in *equivalents of solute per liter of solution*.

Determination of Conductance. The equivalent conductance of an electrolyte is always calculated from specific conductances through equation (21). Consequently the problem of obtaining this quantity reduces itself to a determination of the specific conductance of the electrolyte, and this, in turn, to a measurement of the resistance of the solution and use of equation (20). For measuring resistances of electrolytic solutions the Wheatstone bridge method is employed, a schematic diagram of which

Fig. 6. Principle of Direct Current Wheatstone Bridge

is shown in Fig. 6. R_x, the unknown resistance whose value is to be determined, is placed in one arm of the bridge, a variable known resistance R_s in the other. AB is a uniform slide wire across which moves a contact point C. To balance the bridge the contact is moved

along this resistance until no current from the battery E flows through the galvanometer G, and hence no deflection of it is observed. When this condition is reached, R_s, the resistance from A to C, R_1, and that from C to B, R_2, are read. R_x is then calculated from these according to the following considerations.

When the current I from the battery reaches point B it divides into two parallel paths and into the two currents, I_1 and I_2. These currents lead to potential drops across the resistances they traverse. The purpose of balancing the bridge is to find a point along AB, namely C, such that the potential drop from B to C is equal to that from B to D. When this point is found, D and C are at the same potential and no current flows through the branch containing the galvanometer. The condition for bridge balance is, then,

$$R_s I_1 = R_2 I_2$$

But, when these IR drops are equal, those from D to A and from C to A must also be equal. Hence,

$$R_x I_1 = R_1 I_2$$

Dividing now the second of these equalities by the first, we obtain

$$\frac{R_x}{R_s} = \frac{R_1}{R_2}$$

and therefore,
$$R_x = R_s \left(\frac{R_1}{R_2}\right) \tag{22}$$

By reading R_s, R_1, and R_2 for the bridge balance, R_x can be calculated from equation (22).

Although the principle of the Wheatstone bridge as just given remains the same, several modifications in technique are necessary before resistances of electrolytic solutions can be measured by this method. In the first place, direct current cannot be used, as it would cause electrolysis and concentration changes at the electrodes. To avoid these alternating current is employed usually at a frequency of 1000 cycles per second, and supplied by either a vibrating tuning fork or a vacuum tube oscillator. This current should be close to a pure sine wave in form, so that whatever electrolytic action may tend to take place on one-half of the cycle will be reversed on the other. Since a current of this frequency is within the range of the human ear, the galvanometer can be replaced by a set of earphones. Passage of the current through these produces a buzzing sound which decreases in intensity as balance is approached, and is a minimum when the balance point is reached. Theoretically the sound should be zero at balance, but due to capacitance introduced by

the cell, such an ideal state is not attained. However, by placing a variable condenser across the standard resistance, it is possible to sharpen the balancing by adjusting the condenser to the capacitance of the cell, thus neutralizing the two to a degree. With these modifications the Wheatstone bridge assembly for electrolytic conductivity measurements takes on the form shown in Fig. 7. Here again R_s is the standard resistance, X is the conductivity cell, S is the variable capacitance, H are the headphones, and T is the source of the alternating current.

The cells employed for conductance work are of various types and shapes, depending on the purpose and on the accuracy required. They
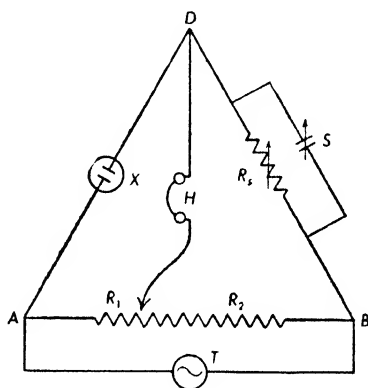
Fig. 7. Assembly for Determination of Electrolytic Conductance

are constructed of glass, with electrodes of either platinum or gold. To overcome imperfections in the current and other effects at the electrodes, the latter are coated electrolytically from a solution of chloroplatinic acid with a thin layer of finely divided platinum, called *platinum black* because of its color. The distance apart the electrodes are placed in a cell is determined by the conductance of the solution to be measured.
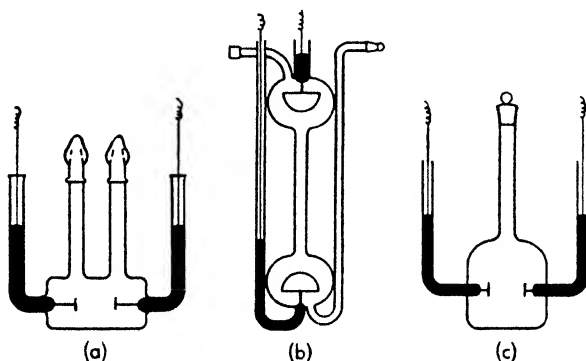
Fig. 8. Types of Conductivity Cells

For solutions of high conductance the electrodes are widely spaced, while for poorly conducting solutions the electrodes are mounted near each other. In Fig. 8 are shown several cells commonly encountered in the laboratory. Types (a) and (b) are used in research, while type (c) is frequently used for less precise work.

The Cell Constant. According to equation (20), the specific conductance of any electrolytic conductor is given by

$$L_s = \left(\frac{l}{A}\right) L = \left(\frac{l}{A}\right)\frac{1}{R} \tag{23}$$

and hence before L_s can be calculated from the measured resistance R the ratio (l/A) for the particular cell used is required. For any given cell this ratio of distance between electrodes to their cross-sectional area is a fixed quantity, and is therefore called the *cell constant K*. To obtain the value of the cell constant it is not necessary to determine l and A. Instead, a solution of known L_s is placed in the cell, the resistance is measured, and K is calculated. Once K is available, the specific conductance of any other solution whose resistance is measured in the same cell follows from

$$L_s = \frac{K}{R} \tag{24}$$

To determine cell constants either 1 demal, 0.1 demal, or 0.01 demal solutions of potassium chloride are used. A demal solution is a solution containing 1 gram mole of salt per *cubic decimeter* of solution at 0° C, or 76.6276 grams of potassium chloride in 1000 grams of water, both weighed in air. The 0.1 and 0.01 demal solutions contain, respectively, 7.47896 and 0.74625 grams of potassium chloride per 1000 grams of water. The conductances of these solutions have been measured with great accuracy in cells with electrodes of definite area and definite spacing, and their specific conductances are well known. The most concentrated solution is used only for cells with electrodes far apart, the other two for electrodes close together or intermediate in spacing. Table 3 gives the specific conductances of these solutions at several temperatures as determined by Jones and Bradshaw.[1]

TABLE 3

SPECIFIC CONDUCTANCES OF KCl SOLUTIONS
(In mhos)

Demal Conc.	Grams KCl per 1000 g H$_2$O (in air)	Grams KCl per 1000 g Solution (in air)	Specific Conductance		
			0° C	18° C	25° C
0.01	0.74625	0.74526	0.00077364	0.00122052	0.00140877
0.10	7.47896	7.41913	0.0071379	0.0111667	0.0128560
1.00	76.6276	71.1352	0.065176	0.097838	0.111342

[1] Jones and Bradshaw, J. Am. Chem. Soc., **55**, 1780 (1933).

The calculation of cell constants, specific, and equivalent conductances from experimental data may be illustrated with the following example. In a particular cell a 0.01 demal solution of potassium chloride gave a resistance of 150.00 ohms at 25°, while a 0.01 N solution of hydrochloric acid gave a resistance of 51.40 ohms at the same temperature. At 25° C the specific conductance of 0.01 demal potassium chloride is 0.0014088, and hence the cell constant for the given cell is

$$K = L_sR = 0.0014088 \times 150$$
$$= 0.21132$$

From this K and the resistance of the hydrochloric acid solution the specific conductance of the latter follows as

$$L_s = \frac{K}{R} = \frac{0.21132}{51.40}$$
$$= 0.004111 \text{ mhos}$$

and the equivalent conductance as

$$\Lambda = \frac{L_s1000}{C}$$
$$= \frac{0.004111 \times 1000}{0.01}$$
$$= 411.1 \text{ mhos/equivalent}$$

Variation of Conductance with Concentration. Both the specific and equivalent conductances of a solution vary considerably with concentration. For strong electrolytes at moderate concentrations, i.e., up to several equivalents per liter, the specific conductance increases sharply with increase in concentration. In contrast to these rapid changes, the specific conductances of weak electrolytes start at lower values in dilute solutions and rise much more gradually with concentration. Thus, in passing from 0.001 to 0.05 N solutions the specific conductance at 25° C for hydrochloric acid changes from 4.21×10^{-4} to 199.5×10^{-4} mhos, that of acetic acid from 0.49×10^{-4} to only 3.70×10^{-4} mhos. In both instances the increase in conductance with concentration is to be ascribed to the increase in the number of ions per unit volume of solution. In strong electrolytes the number of ions per cubic centimeter increases in proportion to the concentration. In weak electrolytes, however, the increase cannot be quite so large because of the changing partial ionization of the solute, and consequently the conductance does not go up so rapidly as in strong electrolytes.

Unlike the specific conductance, the equivalent conductance Λ of both strong and weak electrolytes increases *with dilution*. The reason for this is that the decrease in specific conductance is more than compensated by the increase in the value of $1/C$ on dilution, and hence the value of Λ goes up. The manner in which Λ varies with concentration can be judged from Table 4, and from the plot of Λ vs. $\sqrt{C}$ given in Fig. 9. Why the $\sqrt{C}$ plot is used will be explained later. It may be seen from these

TABLE 4

EQUIVALENT CONDUCTANCES OF ELECTROLYTES IN AQUEOUS SOLUTION AT 25° C [1]

C Equiv./liter	NaCl	KCl	HCl	$AgNO_3$	HIO_3	$NaC_2H_3O_2$	$\frac{1}{2}H_2SO_4$	$\frac{1}{2}Na_2SO_4$	$\frac{1}{2}BaCl_2$	$\frac{1}{2}CuSO_4$	$\frac{1}{3}LaCl_3$	$HC_2H_3O_2$	NH_4OH
0.0000	126.45	149.86	426.16	133.36	—	91.0	429.6	129.9	139.98	—	145.9	390.7	271.4
0.0001	—	—	—	—	—	—	—	—	—	—	—	134.7	93
0.0005	124.50	147.81	422.74	131.36	386.3	89.2	413.1	125.74	135.96	—	139.6	67.7	47
0.001	123.74	146.95	421.36	130.51	383.9	88.5	399.5	124.15	134.34	115.2	137.0	49.2	34
0.005	120.65	143.55	415.80	127.20	370.9	85.72	364.9	117.15	128.02	97.5	127.5	22.9	16
0.01	118.51	141.27	412.00	124.76	359.7	83.76	336.4	112.44	123.94	83.3	121.8	16.3	11.3
0.02	115.76	138.34	407.24	121.41	343.0	81.24	308.0	106.78	119.09	72.2	115.3	11.6	8.0
0.05	111.06	133.37	399.09	115.24	310.7	76.92	272.6	97.75	111.48	58.8	106.2	7.4	5.1
0.10	106.74	128.96	391.32	109.14	278.3	72.80	250.8	89.98	105.19	50.5	99.1	—	3.6
0.20	101.6	123.9	379.6	101.8	242.2	67.7	234.3	81.5	98.6	43.5	—	—	—
0.50	93.3	117.2	359.2	—	219.5	58.6	222.5	—	88.8	35.1	—	—	—
1.00	—	111.9	332.8	—	—	49.1	—	—	80.5	29.3	—	—	—

[1] Taken mostly from MacInnes, "Principles of Electrochemistry," Reinhold Publishing Corporation, New York, 1939, p. 339.

that the Λ's for strong and weak electrolytes behave differently on decrease of concentration. On dilution of a strong electrolyte Λ rapidly approaches already in 0.001 or 0.0001 N solutions a value close to the limiting value of the conductance at zero concentration, Λ_0. On the other hand, although the equivalent conductance of weak electrolytes increases rapidly on dilution, at the concentrations mentioned it is still very far from its limiting value. For instance, Λ at 25° C for 0.001 N sodium chloride is 123.7 as against a Λ_0 of 126.5. At the same concentration and temperature Λ for acetic acid is only 49.2, while Λ_0 for this substance is 390.7. Because of this fundamental difference in the behavior of the equivalent conductances of strong and weak electrolytes on dilution, quite diverse procedures must be used for obtaining their limiting equivalent conductances.
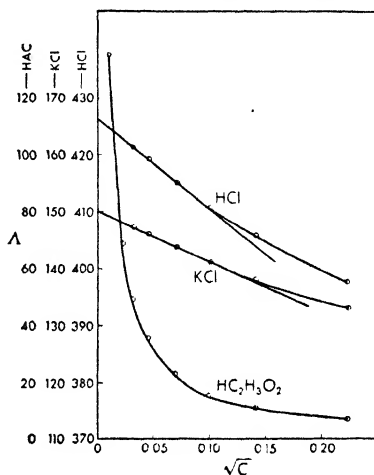
Fig. 9. Plot of Λ vs. $\sqrt{C}$ for Strong and Weak Electrolytes

Equivalent Conductances of Strong Electrolytes at Infinite Dilution. Attention was directed to the fact that on reduction of the concentration of a solution to zero the equivalent conductance approaches a constant limit. This limiting value of the equivalent conductance is known as the *equivalent conductance at infinite dilution*, and is designated by the symbol Λ_0, the subscript indicating zero concentration. The Λ_0 value of an electrolyte is not the same as the equivalent conductance of the pure solvent, but is a characteristic of the electrolyte in question. As this quantity is of great import, its determination is a problem of consequence, and requires some consideration.

Kohlrausch was the first to point out that when Λ for *strong electrolytes* is plotted against $\sqrt{C}$ the curve approaches linearity in dilute solutions, i.e., in dilute solutions the variation of Λ with concentration can be represented by the equation

$$\Lambda = \Lambda_0 - b\sqrt{C} \qquad (25)$$

where b is a constant. The truth of this finding may be seen from the plots for hydrochloric acid and potassium chloride in Fig. 9. Consequently, to obtain Λ_0 of such electrolytes the curve may be extrapolated to $\sqrt{C} = 0$ and the intercept read, or the slope of the linear portion of the curve may be obtained from the plot, and Λ_0 solved for from equa-

tion (25). Some of these Λ_0 values are given in the first horizontal column of Table 4.

Equivalent Conductances of Weak Electrolytes at Infinite Dilution. The method employed for evaluation of Λ_0's of strong electrolytes cannot be used for weak electrolytes. As may be seen from Fig. 9, the plot of Λ vs. $\sqrt{C}$ for a weak electrolyte, acetic acid, does not approach linearity in solutions as dilute as 0.0001 N; instead, Λ exhibits a very rapid increase with decrease in concentration. Again, it is not practicable to carry the measurements to concentrations much lower than 0.0001 N, for in such dilute solutions the conductance of the water becomes an appreciable part of the total conductance. Although the conductance of the water may be subtracted from the specific conductance of the solution to yield that of the electrolyte, such corrections are not always satisfactory and introduce considerable uncertainty into the final result.

Equivalent conductances at infinite dilution for weak electrolytes are evaluated by application of *Kohlrausch's law of independent migration of ions.* This law states that at infinite dilution, where dissociation for all electrolytes is complete and where all interionic affects disappear, each ion can be considered to migrate independently of its co-ion, and thus to contribute its definite share to the total equivalent conductance of the electrolyte. Further, if this be the case, the conductance due to a particular ion at infinite dilution should depend only on the nature of the ion, and not at all on the co-ion with which it is associated. In other words, the law of independent migration of ions states that Λ_0 of any electrolyte is the sum of the equivalent conductances of the ions composing it, provided, of course, that the solvent and temperature are the same. The evidence for the validity of this statement is given in Table 5. According to this law the difference between the Λ_0's of electrolytes containing a common ion should be a constant equal to the difference in the

TABLE 5

KOHLRAUSCH'S LAW OF INDEPENDENT MIGRATION OF IONS

Electrolyte	Λ_0 (25° C)	Difference	Electrolyte	Λ_0 (25° C)	Difference
KCl	149.9		HCl	426.2	
LiCl	115.0	34.9	HNO$_3$	421.3	4.9
KNO$_3$	145.0		KCl	149.9	
LiNO$_3$	110.1	34.9	KNO$_3$	145.0	4.9
KOH	271.5		LiCl	115.0	
LiOH	236.7	34.8	LiNO$_3$	110.1	4.9

equivalent conductances of the ions not in common. In line with this requirement we see that, irrespective of the nature of the co-ion, the difference between the conductances of K^+ and Li^+ is constant and equal to 34.9, and the same is true for the difference in conductance between the Cl^- and NO_3^- ions. The law has been tested on other ions as well, and with the same results.

From the law of independent migration of ions it follows that Λ_0 for any electrolyte may be written as

$$\Lambda_0 = l_+^0 + l_-^0 \tag{26}$$

where l_+^0 and l_-^0 are the *equivalent ionic conductances at infinite dilution* of the cation and anion respectively. Moreover, since the fraction of the total current carried by any ion is given by its transport number, this number must represent as well the fraction of the total conductance due to the ion. Consequently l_+^0 and l_-^0 are also related to Λ_0 by the relations

$$l_+^0 = t_+^0 \Lambda_0 \tag{27}$$

$$l_-^0 = t_-^0 \Lambda_0 \tag{28}$$

t_+^0 and t_-^0 being the transference numbers at infinite dilution as obtained by extrapolation. These equations permit ready calculation of the limiting ionic conductances from transference numbers and Λ_0 values of strong electrolytes. For example, Λ_0 for hydrochloric acid at 25° C is 426.16, while the transport number of the hydrogen ion extrapolated to zero concentration is 0.821. Therefore,

$$l_{H^+}^0 = 0.821 \times 426.16 = 349.9$$
$$l_{Cl^-}^0 = 0.179 \times 426.16 = 76.3$$

By applying this procedure to other electrolytes, the equivalent ionic conductances of other ions have similarly been evaluated. These are summarized in Table 6. The ionic conductances of ions of weak acids or bases were deduced from the Λ_0's of their salts, which are strong electrolytes.

Through equation (26) Λ_0 values of *strong and weak electrolytes* follow from Table 6 on addition of the appropriate ionic conductances of the cation and anion. Thus we find for Λ_0 of acetic acid, a weak electrolyte,

$$\Lambda_{0(HAc)} = l_{H^+}^0 + l_{Ac^-}^0$$
$$= 349.8 + 40.9$$
$$= 390.7 \text{ mhos}$$

In a similar manner Λ_0 for the strong electrolyte silver chloride is

$$\Lambda_0 = l_{Ag^+}^0 + l_{Cl^-}^0$$
$$= 61.92 + 76.34$$
$$= 138.26 \text{ mhos}$$

TABLE 6

EQUIVALENT IONIC CONDUCTANCES AT INFINITE DILUTION [1]
(25° C)

Cations	l_+^0	Anions	l_-^0
K^+	73.52	Cl^-	76.34
Na^+	50.11	Br^-	78.4
Li^+	38.69	I^-	76.8
NH_4^+	73.4	NO_3^-	71.44
H^+	349.82	HCO_3^-	44.48
Ag^+	61.92	OH^-	198
Tl^+	74.7	Acetate$^-$	40.9
$1/2\ Ca^{++}$	59.50	Chloracetate$^-$	39.7
$1/2\ Ba^{++}$	63.64	Propionate$^-$	35.81
$1/2\ Sr^{++}$	59.46	ClO_4^-	68.0
$1/2\ Mg^{++}$	53.06	$1/2\ SO_4^{--}$	79.8
$1/3\ La^{+++}$	69.6	$1/3\ Fe(CN)_6^{---}$	101.0
		$1/4\ Fe(CN)_6^{----}$	110.5

In arriving at the equivalent ionic conductances given in Table 6, and from these at the Λ_0's of electrolytes, transference numbers were used. However, it is possible to obtain Λ_0's of electrolytes by direct addition and subtraction of appropriate Λ_0 values without using transport numbers. Thus, if Λ_0 for hydrochloric acid is added to that for sodium acetate, and that for sodium chloride subtracted, the result is Λ_0 for acetic acid. For

$$\Lambda_{0(NaAc)} + \Lambda_{0(HCl)} - \Lambda_{0(NaCl)} = l_{Na^+}^0 + l_{Ac^-}^0 + l_{H^+}^0 + l_{Cl^-}^0 - l_{Na^+}^0 + l_{Cl^-}^0$$
$$= l_{H^+}^0 + l_{Ac^-}^0$$
$$= \Lambda_{0(HAc)}$$

In a like manner, Λ_0 of nitric acid follows from the Λ_0's of potassium nitrate, potassium chloride, and hydrochloric acid. This method of calculating Λ_0 is particularly valuable with weak electrolytes for whose salt solutions transport numbers are not available, and neither are the equivalent ionic conductances at infinite dilution for both ions. From the example cited it is evident that all that is necessary to obtain Λ_0 for a weak acid for which no data are extant is Λ_0 of its sodium salt. Since the salt is a strong electrolyte, its Λ_0 can be evaluated from the measured equivalent conductances without any special difficulty. By combining this Λ_0 with that of hydrochloric acid and sodium chloride, as shown above, the Λ_0 for the acid follows immediately. Similarly, Λ_0 for a weak base like ammonium hydroxide can be calculated from Λ_0 of ammonium chloride, sodium hydroxide, and sodium chloride.

[1] MacInnes, op. cit., p. 342.

Effect of Other Factors on Conductance. The conductance of all electrolytes increases with temperature. The variation of Λ_0 with temperature can be represented by the equation

$$\Lambda_{0(t)} = \Lambda_{0(25° C)}[1 + \beta(t - 25)] \tag{29}$$

where $\Lambda_{0(t)}$ is the limiting equivalent conductance at $t°$ C, $\Lambda_{0(25° C)}$ that at 25° C, and β a constant. β for salts is usually 0.022 to 0.025, indicating that Λ_0 increases 2.2 to 2.5 per cent per degree. For acids this constant is somewhat smaller, being 0.016 to 0.019. Similar behavior is exhibited by the equivalent conductances of strong electrolytes in finite concentrations. However, with weak electrolytes the variation of Λ with temperature is not so regular, for in these not only do the velocities of the ions and the interionic forces change, but also the degree of dissociation.

The conductance behavior observed in nonaqueous solvents depends pretty much on the dielectric constant of the medium. The dielectric constant of water is high, 78.6 at 25° C, whereas that of most other solvents is considerably lower. Thus the dielectric constants D at 25° C for methyl alcohol, ethyl alcohol, and dioxane are, respectively, 31.5, 24.3, and 2.2. As the dielectric constant of a solvent is lowered, the conductance of an electrolyte in the medium also decreases. Beyond this conductance drop, and beyond the fact that some electrolytes that are strong in water may be weak in other solvents, the conductance behavior of these substances is not very different in nonaqueous solvents from that in water, provided the dielectric constant is above about 25. For example, the halides and nitrates of alkali metals, the thiocyanates of the alkali and alkaline earth metals, and the tetraalkyl ammonium salts are strong electrolytes in ethyl and methyl alcohols, and behave pretty much as do strong electrolytes in water. Again, picric acid, various substituted acetic acids, phenols, and cresols are weak electrolytes in methyl alcohol, as are also hydrochloric, hydrobromic, hydriodic, and picric acids in ethyl alcohol. In these solvents the substances mentioned exhibit conductance curves typical of this class. The influence of the solvent on the conductance can be judged from Table 7, where the Λ_0 values at 25° C for several electrolytes obtained in water and methyl and ethyl alcohol are compared.

In solvents of dielectric constant less than 25, the dependence of the equivalent conductance on the concentration becomes complex. Plots of log Λ vs. log C, instead of being linear or slightly curved as they are in solvents of higher dielectric constant, contain minima which appear at lower concentrations the lower the dielectric constant. To explain these minima and the curves in general it has been suggested [1] that in these

[1] Kraus and Fuoss, *J. Am. Chem. Soc.* **55**, 21 (1933); **55**, 476, 1019, 2387 (1933); **57**, 1 (1935).

TABLE 7

Λ_0 FOR SOME ELECTROLYTES IN VARIOUS SOLVENTS
(25° C)

Electrolyte	Solvent		
	$H_2O (D = 78.6)$	$CH_3OH (D = 31.5)$	$C_2H_5OH (D = 24.3)$
HCl	426.16	—	81.8
LiCl	115.03	90.9	39.2
NaCl	126.45	69.9	42.5
KCl	149.86	105.0	—
LiNO₃	110.13	100.2	42.7
KNO₃	118.00	114.5	—

solvents ions exhibit a tendency to associate into complexes such as A^+B^-, $A^+B^-A^+$, and $B^-A^+B^-$, which decrease the number of ions available to carry current, and hence the conductance. These theories seem to account, at least in part, for the observed phenomena.

The Interionic Attraction Theory of Conductance. The decrease in the equivalent conductance with increase in concentration for weak electrolytes can be explained as due essentially to a decrease in the degree of ionization. However, from everything that has been said heretofore it is evident that such an explanation cannot apply to strong electrolytes, for these, at least in the more dilute solutions, are completely dissociated. Consequently, to account for the variation of Λ with concentration in strong electrolytes some other explanation must be sought, and this is found in the Debye-Hückel-Onsager theory of conductance.

According to the Debye-Hückel theory of interionic attraction, as developed briefly in Chapter VII, each ion in solution is surrounded by an atmosphere of other ions whose net charge is on an average opposite to that of the central ion. When the ions have no external force applied upon them, this atmosphere is spherically and symmetrically distributed about the ion. However, when an external force is imposed on the solution, as when a potential is applied across two electrodes immersed in the solution during conductance, the ions are set in motion, and as a consequence certain effects and changes in the ionic atmosphere arise which result in a decrease in the speeds of the ions. Debye and Hückel[1] first pointed out that these effects are twofold, namely, (1) the *relaxation of the ionic atmosphere* due to an applied potential, and (2) the *electrophoretic effect*. The first of these arises from the fact that any central ion and its atmosphere are oppositely charged, i.e., when the central ion is positively charged the atmosphere is negative, and vice versa. Because of this

[1] Debye and Hückel, Physik. Zeit., **24**, 185, 305 (1923).

difference in sign of the atmosphere and central ion, a potential applied across the combination will tend to move the central ion in one direction, the atmosphere in the other. Thus a central positive ion will tend to move toward the cathode while its ionic atmosphere will tend toward the anode. The symmetry of the atmosphere about an ion is destroyed by these opposing tendencies, and the atmosphere becomes distorted. In this state the force exerted by the atmosphere on the ion is no longer uniform in all directions, but is greater *behind* the ion than in front of it. Consequently the ion experiences a retarding force opposite to the direction of its motion, and the ion is slowed down by these interionic attractions.

The electrophoretic effect arises from the fact that an ion, in moving through the solution, does not travel through a stationary medium, but through one that moves in a direction opposite to that of the ion. Ions are generally solvated, and when these move, they carry with them solvent. Any positive ion migrating toward the cathode has, then, to thread its way through medium moving with the negative ions toward the anode. Similarly negative ions have to migrate through molecules of solvent carried by positive ions in the opposite direction. These countercurrents make it more difficult for the ion to move through the solution, and thus slow it down in the same way as swimming against the current in a river would slow down a swimmer.

Debye and Hückel showed that both of these retarding effects on an ion produce a decrease in equivalent conductance dependent on the concentration. Their mathematical analysis of the problem was subsequently extended by Lars Onsager [1] to include not only the relaxation of the ionic atmosphere and the electrophoretic effects, but also the natural Brownian movement of the ions. As a result of his treatment Onsager finally arrived at the following equation for the dependence of the equivalent conductance of a binary strong electrolyte on the concentration,

$$\Lambda = \Lambda_0 - \left[\frac{0.9834 \times 10^6}{(DT)^{3/2}} w\Lambda_0 + \frac{28.94(z_+ + z_-)}{\eta(DT)^{1/2}} \right] \sqrt{(z_+ + z_-)C} \quad (30)$$

where
$$w = z_+z_- \left(\frac{2q}{1+\sqrt{q}} \right)$$

$$\bar{q} = \frac{z_+z_-\Lambda_0}{(z_+ + z_-)(z_+l^0_- + z_-l^0_+)}$$

In this equation Λ, Λ_0, C, l^0_+, and l^0_- have the same significance as before, z_+ and z_- are the charges of the two ions, T is the absolute temperature, and D and η are the dielectric constant and viscosity of the solvent, respectively. Because of simplifications in its derivation, this equation

[1] Onsager, Physik. Zeit., **27**, 388 (1926); **28**, 277 (1927).

can be applied only to very dilute solutions, and must be considered, therefore, as essentially a limiting equation for conductance. *For the special case of 1–1 electrolytes in water at 25° C*, for which $z_+ = z_- = 1$, $D = 78.55$, and $\eta = 0.008949$ poise, equation (30) reduces to

$$\Lambda = \Lambda_0 - [\theta\Lambda_0 + \sigma] \sqrt{C} \tag{31}$$

where θ and σ are constants with values $\theta = 0.2273$, and $\sigma = 59.78$. Since the quantity in brackets is a constant, equation (31) is identical in form with equation (25) proposed by Kohlrausch on empirical grounds for strong electrolytes in dilute solution.

From preceding discussion of Kohlrausch's equation (25) it is evident that the theoretical equation (31) has the correct form for the dependence of Λ on $\sqrt{C}$. The only question remaining is whether the experimental slope is in accord with that predicted by the Onsager equation, namely, slope $= [\theta\Lambda_0 + \sigma]$. Exhaustive tests by Shedlovsky, MacInnes, and others indicate that in very dilute solutions the Onsager equation is in accord with observation not only for 1–1 strong electrolytes, but also for electrolytes of higher valence types, such as calcium chloride and lanthanum chloride. The concordance obtained may be judged from the fact that for potassium chloride Λ_0 found by graphical extrapolation is 149.86, while that calculated by equation (31) from the experimental data is 149.98.

Equation (31) is applicable not only to aqueous solutions, but to strong electrolytes in other solvents as well. However, the concordance usually is not so good as in water. In some instances the slopes predicted for the Λ vs. $\sqrt{C}$ plots deviate from the experimental quite appreciably, and may be off as much as 100 per cent or over, as 126 per cent for silver nitrate in ethyl alcohol at 25° C.

Absolute Velocities of Ions. The absolute velocity in centimeters per second with which any ion moves through a solution depends on the nature of the ion, the concentration of the solution, the temperature, and the applied potential drop per centimeter of conducting path. Such velocities can be measured directly by application of the moving boundary method. However, the velocities, or *mobilities* as they are generally called, may also be calculated from conductance measurements, and to this end we turn now our attention.

For this purpose consider a pair of parallel electrodes 1 cm apart, such as those illustrated in Fig. 2, across which is applied a potential difference $\mathcal{E}$, and between which is contained a volume of solution containing 1 mole of an electrolyte A_xB_y. For the sake of generality, consider this electrolyte to be ionized to an extent α according to the equation

$$\begin{array}{cccc} A_xB_y & = xA^{z_+} + yB^{z_-} \\ 1(1-\alpha) & x\alpha & y\alpha \end{array} \tag{32}$$
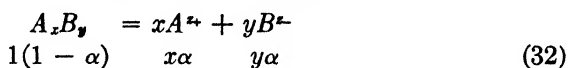
where z_+ and z_- are the valences of the ions, while x and y are the numbers of these obtained from one molecule of $A_x B_y$. The concentrations of the various species are, then, as indicated in equation (32), and the number of positive ions present between the two electrodes, n_+, is $n_+ = x\alpha N$, where $N =$ Avogadro's number. Now, according to equation (10) of this chapter, the current I flowing between the plates is given by

$$I = n_+ z_+ e(v_+ + v_-)$$

Substituting for n_+ its equivalent $x\alpha N$, we have

$$I = x\alpha N z_+ e(v_+ + v_-)$$
$$= \alpha \mathcal{F}(xz_+)(v_+ + v_-) \qquad (33)$$

where $\mathcal{F}$, the value of the faraday, was set equal to Ne.

One mole of $A_x B_y$ corresponds to xz_+ equivalents of substance, i.e., the number of positive ions per molecule multiplied by the valence. Therefore, if the concentration of the solution is C equivalents per 1000 cc of solution, the volume per equivalent is $1000/C$, and per mole $(xz_+ 1000)/C$ cc. Since the electrodes are 1 cm apart, the conductance of 1 cc of solution is the specific conductance L_s, and hence the total conductance L of the solution between the plates is L_s times the total volume, or

$$L = \left(\frac{L_s\,1000}{C}\right)xz_+ = \Lambda xz_+ \qquad (34)$$

Again, by Ohm's law,

$$I = \frac{\mathcal{E}}{R} = \mathcal{E}L$$

and hence, on substitution of I from equation (33) and L from equation (34), we find

$$\alpha \mathcal{F}(xz_+)(v_+ + v_-) = \mathcal{E}\Lambda xz_+$$
$$\mathcal{E}\Lambda = \alpha \mathcal{F}(v_+ + v_-) \qquad (35)$$

For the special case of a potential drop of *1 volt per centimeter*, $\mathcal{E} = 1$, and equation (35) reduces to

$$\Lambda = \alpha \mathcal{F}(v_+ + v_-) \qquad (36)$$

According to Kohlrausch's law of independent migration of ions,

$$\Lambda_0 = l_+^0 + l_-^0 \qquad (37a)$$

For infinite dilution $\alpha = 1$, $\Lambda = \Lambda_0$, $v_+ = v_+^0$, $v_- = v_-^0$, and equation (36) becomes

$$\Lambda_0 = \mathcal{F}(v_+^0 + v_-^0)$$
$$= \mathcal{F}v_+^0 + \mathcal{F}v_-^0 \qquad (37b)$$

A comparison of (37a) and (37b) readily reveals that

$$\mathscr{J}v_+^0 = l_+^0$$

$$v_+^0 = \frac{l_+^0}{\mathscr{J}} \tag{38}$$

and

$$\mathscr{J}v_-^0 = l_-^0$$

$$v_-^0 = \frac{l_-^0}{\mathscr{J}} \tag{39}$$

i.e., *at infinite dilution the velocity of any ion in centimeters per second, under a potential drop of 1 volt per centimeter, is given by the limiting equivalent conductance of the ion divided by the value of the faraday.*

The mobilities of a number of ions calculated with the aid of equations (38) and (39) from the equivalent ionic conductances are given in Table 8. It will be observed that the velocities are unusually low. Further, with the exception of the hydrogen and hydroxyl ions, the speeds are not very different for the various ions. The hydrogen and hydroxyl ions are unique in their high mobility, and this fact is the basis of several applications of conductance measurements.

If equation (36) is rewritten

$$\Lambda = \alpha \mathscr{J}v_+ + \alpha \mathscr{J}v_- \tag{36a}$$

the new form suggests that Λ at *finite* concentration can also be represented by

$$\Lambda = l_+ + l_- \tag{40}$$

where l_+ and l_- are now the equivalent ionic conductances for the concentration to which Λ corresponds. Analogously to equations (27) and (28) these equivalent conductances of the ions can be evaluated from Λ and the transport numbers by the relations

$$l_+ = t_+\Lambda \tag{41}$$

$$l_- = t_-\Lambda \tag{42}$$

By comparison of equations (36) and (40), the mobilities of the ions at any given concentration follow as

$$v_+ = \frac{l_+}{\alpha \mathscr{J}} \tag{43}$$

$$v_- = \frac{l_-}{\alpha \mathscr{J}} \tag{44}$$

for weak electrolytes, and

$$v_+ = \frac{l_+}{\mathscr{J}} = \frac{t_+\Lambda}{\mathscr{J}} \tag{45}$$

$$v_- = \frac{l_-}{\mathscr{J}} = \frac{t_-\Lambda}{\mathscr{J}} \tag{46}$$

for strong electrolytes, since these are completely ionized.

TABLE 8

ABSOLUTE VELOCITIES OF IONS AT 25° C
(Cm/second/volt/centimeter)

Cation	v_+^0	Anion	v_-^0
K^+	0.000762	Cl^-	0.000791
Na^+	0.000520	Br^-	0.000812
Li^+	0.000388	I^-	0.000796
NH_4^+	0.000760	NO_3^-	0.000740
H^+	0.003620	HCO_3^-	0.000461
Ag^+	0.000642	OH^-	0.002050
Tl^+	0.000774	$C_2H_3O_2^-$	0.000424
Ca^{++}	0.000616	$C_3H_5O_2^-$	0.000411
Ba^{++}	0.000659	ClO_4^-	0.000705
Sr^{++}	0.000616	SO_4^{--}	0.000827
Mg^{++}	0.000550	$Fe(CN)_6^{---}$	0.001040
La^{+++}	0.000721	$Fe(CN)_6^{----}$	0.001140

Degree of Ionization and Conductance. In developing his theory of electrolytic dissociation, Arrhenius ascribed the decrease in Λ with increase in concentration entirely to variation of the degree of dissociation of the electrolyte with concentration. If this is correct, it can readily be shown that the degree of dissociation α must be given by the ratio Λ/Λ_0. From equations (36) and (37b), Λ/Λ_0 for any electrolyte follows as

$$\frac{\Lambda}{\Lambda_0} = \frac{\alpha(v_+ + v_-)}{(v_+^0 + v_-^0)} \qquad (47)$$

Assuming, further, that the velocities of the ions at any finite concentration are the same as at infinite dilution, then $(v_+ + v_-) = (v_+^0 + v_-^0)$, and, therefore,

$$\alpha = \frac{\Lambda}{\Lambda_0} \qquad (48)$$

According to this equation the degree of dissociation of any electrolyte is unity at infinite dilution, i.e., the electrolyte is completely dissociated. On the other hand, as the concentration is increased Λ falls, and so does the value of α.

Arrhenius supposed equation (48) to be applicable to both strong and weak electrolytes. However, in view of our present ideas of the complete dissociation of strong electrolytes, Λ/Λ_0 cannot possibly have the significance given it by Arrhenius. Rather, this ratio must be looked upon as a measure of the effect of interionic forces upon the velocities of the ions. In line with this argument is the fact that the Debye-Hückel-Onsager theory of conductance of strong electrolytes can account for the

variation of Λ with concentration in terms of interionic forces only, without the aid of partial dissociation. Again, careful measurements have revealed that transference numbers, and therefore ionic velocities, are not independent of concentration, and consequently the assumption introduced in passing from equation (47) to equation (48) cannot be valid for such electrolytes.

However, applications of the law of mass action to weak electrolytes, to be discussed in the following chapter, have definitely established that the conductance ratio Λ/Λ_0 does represent, within a small correction, the degree of dissociation. In weak electrolytes the variation of Λ with concentration is due to two factors, namely, (1) the partial dissociation of the electrolyte, and (2) the interionic attractions between ions present. Because of the relatively low degree of dissociation of these electrolytes, the concentrations of ions present in any given solution are quite low, and so are the interionic effects. Consequently most of the decrease in equivalent conductance is due to decrease in α, and within a fairly close approximation equation (48) holds true.

In a weak electrolyte such as acetic acid at a concentration C and degree of dissociation α, the concentration of dissociated acid is $C\alpha$. Because of this concentration of dissociated acid interionic forces are present that tend to decrease the value of Λ from Λ_0 to Λ_e, where Λ_e is the equivalent conductance of the completely dissociated acid at a concentration $C\alpha$. Consequently, in order to obtain the degree of dissociation of the electrolyte corrected for the interionic effects it would be more nearly correct to calculate α, not as Λ/Λ_0, but as

$$\alpha = \frac{\Lambda}{\Lambda_e} \tag{49}$$

i.e., as the ratio of the actual equivalent conductance to the equivalent conductance of the same electrolyte when completely dissociated at the same ionic concentration as is present in the solution. The evaluation of Λ_e for use in equation (49) may be illustrated with the following example. At $C = 0.1000$, Λ for acetic at 25° C is 5.201, while $\Lambda_0 = 390.71$. From these data we find provisionally that

$$\alpha = \frac{\Lambda}{\Lambda_0} = \frac{5.201}{390.71} = 0.0133 = 1.33\%$$

and hence $C\alpha = 0.00133$ equivalent per liter. Now, on the supposition that acetic acid is completely dissociated, and on the assumption that Kohlrausch's law is applicable, it can be calculated from the Λ's of sodium acetate, sodium chloride, and hydrochloric acid, all at $C = 0.00133$, that

PROBLEMS

1. A potential of 110 volts d-c is applied to the terminals of an electric lamp, and a current of 2 amp is found to flow through the lamp. (a) What is the resistance of the lamp, and (b) how many calories of heat are dissipated per hour?

2. A constant direct current flows through an iodine coulometer for a period of 2 hr. At the end of this time 25.2 cc of 0.08 molar $Na_2S_2O_3$ are required to react with the liberated I_2. What was the current passing through the coulometer?
   Ans. 0.027 amp

3. The platinum crucible used in a silver coulometer gains 0.500 g in a certain electrolysis. What would be the gain in weight of a copper cathode in a cell filled with potassium cuprocyanide [$KCu(CN)_2$] placed in the same circuit?

4. What volume of O_2 would be liberated from an aqueous solution of NaOH by a current of 2 amp flowing for 1 1/2 hr? The temperature is 27° C and the total pressure is 1 atm.   *Ans.* 7138 cc

5. (a) How long would it take a current of 1 amp to reduce completely 80 cc of 0.1 molar $Fe_2(SO_4)_3$ to $FeSO_4$? (b) How many cc of 0.1 molar $K_2Cr_2O_7$ could be reduced to chromic sulfate, $Cr_2(SO_4)_3$, by the same quantity of electricity?

6. What quantity of electricity would be required to reduce 10 g of nitrobenzene completely to aniline? If the potential drop across the cell is 2 volts, how much energy, in calories, is consumed in the process?
   Ans. 47,070 coulombs;   22,510 cal

7. A 4 molal solution of $FeCl_3$ is electrolyzed between platinum electrodes. After the electrolysis the cathode portion, weighing 30 g, is 3.15 molal in $FeCl_3$ and 1.00 molal in $FeCl_2$. What are the transport numbers of Fe^{+++} and Cl^- ions?   *Ans.* $t_+ = 0.448$

8. A $AgNO_3$ solution containing 0.00739 g of $AgNO_3$ per gram of H_2O is electrolyzed between silver electrodes. During the experiment 0.078 g of Ag plate out at the cathode. At the end of the experiment the anode portion contains 23.14 g of H_2O and 0.236 g of $AgNO_3$. What are the transport numbers of Ag^+ and NO_3^- ions?

9. The transference numbers of the ions of 1.000 N KCl were determined by the moving boundary method using a solution 0.80 N $BaCl_2$ as the following solution. Using a current of 0.0142 amp the time required for the boundary to sweep through a volume of 0.1205 cc was 1675 sec. What are the transport numbers of K^+ and Cl^- ions?   *Ans.* $t_+ = 0.47$

10. The cathode, center, and anode chambers of an electrolytic cell contain each 10 milliequivalents of HCl in aqueous solution. What will be the number of milliequivalents of HCl in each compartment after the passage of 5 milliequivalents of electricity through the cell? Assume that t_+ is 0.8 and that H_2 is given off at the cathode and Cl_2 at the anode.

11. At 25° C a cell filled with 0.01 demal KCl solution gave a resistance of 484.0 ohms. The following data for NaCl solutions were then taken in the same cell at 25° C:

Normality	Resistance (ohms)
0.0005	10,910
0.0010	5,494
0.0020	2,772
0.0050	1,128.9

(a) Calculate Λ for NaCl at each concentration, and (b) evaluate Λ_0 by plotting Λ against $\sqrt{C}$ and extrapolating to infinite dilution.

12. From the following equivalent conductances at infinite dilution at 18° C, find Λ_0 for NH_4OH at 18° C:

$$Ba(OH)_2 : \Lambda_0 = 228.8$$
$$BaCl_2 : \Lambda_0 = 120.3$$
$$NH_4Cl : \Lambda_0 = 129.8$$

13. At 25° C the equivalent conductance of a 0.02 molar $AgNO_3$ solution is 128.7, while the transport number of Ag^+ is 0.477. Calculate the equivalent ionic conductances and the ionic mobilities of Ag^+ and NO_3^- in a 0.02 molar solution of $AgNO_3$.

14. In measuring the mobility of H^+ ions by the moving boundary procedure, it is observed that the boundary moves a distance of 4.0 cm in 12.52 min. The voltage drop across the cell before the formation of the boundary is 16.0 volts. The distance between electrodes is 9.6 cm. Calculate from these data the mobility and ionic conductance of hydrogen ions.   *Ans.* $v_+ = 0.0032$ cm/sec

15. At 18° C the mobility at infinite dilution of the ammonium ion is 0.00066 cm/second, while that of the chlorate ion is 0.00057 cm/second. Calculate Λ_0 of ammonium chlorate and the transport numbers of the two ions.

16. At 25° C the resistance of a cell filled with 0.01 demal KCl solution is 525 ohms. The resistance of the same cell filled with 0.1 N NH_4OH is 2030 ohms. What is the degree of dissociation of NH_4OH in this solution?   *Ans.* $\alpha = 0.0134$

17. What will be the resistance of the cell used in the preceding problem when it is filled with H_2O having a specific conductance of 2×10^{-6} mhos?

18. The specific conductance at 25° C of a saturated aqueous solution of $SrSO_4$ is 1.482×10^{-4} mhos, while that of the H_2O used is 1.5×10^{-6} mhos. Using the data in Table 6, determine at 25° C the solubility in grams per liter of $SrSO_4$ in water.   *Ans.* 0.0924 g/liter

19. In the titration of 25.0 cc (diluted to 300 cc) of a $NaC_2H_3O_2$ solution with 0.0972 N HCl solution the following data were found:

Volume of HCl Used	Conductance $\times 10^4$
10.0	3.32
15.0	3.38
20.0	3.46
45.0	4.64
50.0	5.85
55.0	7.10

What is the strength of the $NaC_2H_3O_2$ solution in moles per liter?

20. Sketch roughly the course of conductance against volume of reagent curves for the following titrations: (a) C_6H_5OH with NaOH; (b) $CuSO_4$ with NaOH; (c) K_2CrO_4 with $AgNO_3$.

Ionic Equilibria

Although the law of mass action in its exact form demands that activities rather than concentrations must be employed in writing equilibrium constants, we have seen that equilibrium constants for gaseous equilibria and for nonionic reactions in solution can be expressed in terms of concentrations to a fairly good approximation at relatively low pressures and concentrations. However, in dealing with ionic equilibria, with certain aspects of kinetics in solution, and with electromotive force studies, substitution of concentrations for activities is not always possible. For this reason it is essential before approaching these subjects to consider how ionic concentrations may be converted to activities, and how such activities can be evaluated.

Activities and Activity Coefficients of Strong Electrolytes. In order to introduce some definitions commonly employed in dealing with the activities of strong electrolytes, consider an electrolyte A_xB_y which ionizes in solution according to

$$A_xB_y = xA^{z_+} + yB^{z_-} \tag{1}$$

where z_+ and z_- are the charges of the two ions. The activity of the electrolyte as a whole, a_2, is defined in terms of the activities of the two ions a_+ and a_- as

$$a_2 = a_+^x a_-^y \tag{2}$$

If we designate now by $\nu = x + y$ the total number of ions resulting from 1 molecule of the electrolyte, *the geometric mean activity of the electrolyte,* or more simply the *mean activity,* written $a_\pm$, is defined as

$$a_\pm = \sqrt[\nu]{a_2} = \sqrt[\nu]{a_+^x a_-^y} \tag{3}$$

To relate the activities of the ions to their concentrations we may write that

$$a_+ = C_+ f_+ \tag{4a}$$
$$a_- = C_- f_- \tag{4b}$$

where C_+ and C_- are the gram ionic weights per liter of the two ions in solution, while f_+ and f_- are the *activity coefficients* of the two ions. These

activity coefficients are appropriate factors which when multiplied by the concentrations of the respective ions yield their activities. Introducing equations (4a) and (4b) into equation (2), we obtain for a_2

$$a_2 = (C_+f_+)^x(C_-f_-)^y$$
$$= (C_+^x C_-^y)(f_+^x f_-^y) \tag{5}$$

and for the mean activity, from equation (3),

$$a_\pm = \sqrt[\nu]{a_2} = \sqrt[\nu]{(C_+^x C_-^y)(f_+^x f_-^y)}$$
$$= (C_+^x C_-^y)^{1/\nu}(f_+^x f_-^y)^{1/\nu} \tag{6}$$

The factor $(f_+^x f_-^y)^{1/\nu}$ is generally designated as the *mean activity coefficient* of the electrolyte, f; i.e.,

$$f = (f_+^x f_-^y)^{1/\nu} \tag{7}$$

Similarly the factor $(C_+^x C_-^y)^{1/\nu}$ is defined as the *mean molarity* of the electrolyte, $C_\pm$, or

$$C_\pm = (C_+^x C_-^y)^{1/\nu} \tag{8}$$

In terms of the mean molarity and mean activity coefficient, equations (5) and (6) may be written simply as

$$a_\pm = a_2^{1/\nu} = C_\pm f \tag{9}$$
$$a_2 = a_\pm^\nu = (C_\pm f)^\nu \tag{10}$$

Finally, since for any electrolyte of molarity C $C_+ = xC$ and $C_- = yC$, equations (9) and (10) become also

$$a_\pm = a_2^{1/\nu} = [(xC)^x(yC)^y]^{1/\nu}f$$
$$= (x^x y^y)^{1/\nu}Cf \tag{11}$$
$$a_2 = a_\pm^\nu = (x^x y^y)C^\nu f^\nu \tag{12}$$

Equations (11) and (12) are the expressions needed for converting activities to molarities, or vice versa. Although these expressions may appear complicated, they are actually not so when applied to specific cases. Thus for a 1–1 type of electrolyte, such as sodium chloride, of molarity C we have $x = 1$, $y = 1$, $\nu = 2$, and therefore, by equations (11) and (12),

$$a_\pm = (1 \times 1)^{1/2}Cf = Cf$$
$$a_2 = a_\pm^2 = C^2 f^2$$

Again, for an electrolyte of the 2–1 type, such as barium chloride, we obtain $x = 1$, $y = 2$, $\nu = 3$, and hence

$$a_\pm = (1 \times 2^2)^{1/3}Cf$$
$$= \sqrt[3]{4}\ Cf$$
$$a_2 = a_\pm^3 = 4\ C^3 f^3$$

In Table 1 are summarized the relations connecting C, f, $a_\pm$, and a_2 for a number of different types of electrolytes. It will be observed that the expressions resulting from equations (11) and (12) depend on the electrolyte type, and are identical for the 1-1, 2-2, 3-3 types; for the 1-2 and 2-1 types; for the 1-3 and 3-1 types; and for the 2-3 and 3-2 types.

TABLE 1

RELATION OF $a_\pm$ AND a_2 TO C AND f FOR VARIOUS ELECTROLYTES

Electrolyte Type	Example	x	y	ν	$C_\pm$	$a_\pm = C_+ f$	$a_2 = a_\pm^\nu$
1-1	NaCl	1	1	2	C	Cf	$C^2 f^2$
2-2	CuSO$_4$	1	1	2	C	Cf	$C^2 f^2$
3-3	AlPO$_4$	1	1	2	C	Cf	$C^2 f^2$
1-2	Na$_2$SO$_4$	2	1	3	$\sqrt[3]{4}\,C$	$\sqrt[3]{4}\,Cf$	$4\,C^3 f^3$
2-1	BaCl$_2$	1	2	3	$\sqrt[3]{4}\,C$	$\sqrt[3]{4}\,Cf$	$4\,C^3 f^3$
1-3	Na$_3$PO$_4$	3	1	4	$\sqrt[4]{27}\,C$	$\sqrt[4]{27}\,Cf$	$27\,C^4 f^4$
3-1	La(NO$_3$)$_3$	1	3	4	$\sqrt[4]{27}\,C$	$\sqrt[4]{27}\,Cf$	$27\,C^4 f^4$
2-3	Ca$_3$(PO$_4$)$_2$	3	2	5	$\sqrt[5]{108}\,C$	$\sqrt[5]{108}\,Cf$	$108\,C^5 f^5$
3-2	La$_2$(SO$_4$)$_3$	2	3	5	$\sqrt[5]{108}\,C$	$\sqrt[5]{108}\,Cf$	$108\,C^5 f^5$

The definitions for the ionic and mean activity coefficients have been expressed in terms of concentration on a molarity basis, i.e., moles per liter of solution. In electrochemical work quite frequently concentrations are expressed on a molality basis, i.e., moles per 1000 grams of solvent, m. When this is the case the activities of the ions are defined analogously to equation (4) by

$$a_+ = m_+ \gamma_+ \qquad (13a)$$
$$a_- = m_- \gamma_- \qquad (13b)$$

By repeating the above argument with these new definitions, it can readily be shown that now

$$a_\pm = (x^x y^\nu)^{1/\nu} m\gamma \qquad (14)$$
$$a_2 = a_\pm^\nu = (x^x y^\nu) m^\nu \gamma^\nu \qquad (15)$$

where γ is the *mean activity coefficient of the electrolyte* for concentration in molalities, and is given by

$$\gamma = (\gamma_+^x \gamma_-^y)^{1/\nu} \qquad (16)$$

With equations (14) and (15) the expressions for the mean molality, $m_\pm$, i.e.,

$$m_\pm = (m_+^x m_-^y)^{1/\nu} \qquad (17)$$

$a_\pm$, and a_2 in Table 1 have exactly the same form as for molarities, except that C and f are replaced by m and γ. Since C and m are not identical,

neither will f as a rule be equal to γ. In fact, the two activity coefficients are related by the expression

$$f = \gamma\left(\frac{d_0 m}{C}\right) \qquad (18)$$

where d_0 is the density of the pure solvent, and m/C is given by

$$\frac{m}{C} = \frac{1 + 0.001\, mM_2}{d} \qquad (19)$$

Here M_2 is the molecular weight of the electrolyte, while d is the density of the solution. From equations (18) and (19) it can be shown that for dilute solutions f will be essentially equal to γ; however, in more concentrated solutions the two mean activity coefficients will have different values.

Determination of Activity Coefficients. Equations (11) and (12) or (14) and (15) indicate that for conversion of molarities or molalities to activities the mean activity coefficients for various concentrations of an electrolyte must be known. Such mean activity coefficients of electrolyte solutions can be determined from vapor pressure, freezing point lowering, boiling point elevation, osmotic pressure, distribution, solubility, and electromotive force measurements by well worked out and familiar thermodynamic methods. Although in appropriate places in this book the solubility and electromotive force methods for evaluating activity coefficients will be explained, a discussion of the other methods is beyond our scope. For these books on chemical thermodynamics should be consulted.

All evaluations of activity coefficients are made on the supposition that in the infinitely dilute solution $\gamma = 1$ or $f = 1$, depending on which concentration scale is employed; i.e., that

$$\frac{a_\pm}{m_\pm} = \gamma = 1 \text{ as } m \longrightarrow 0 \qquad (20)$$

or that

$$\frac{a_\pm}{C_\pm} = f = 1 \text{ as } C \longrightarrow 0 \qquad (21)$$

These definitions are equivalent to the statement that the activity of an ion is equal to its concentration in the infinitely dilute solution. On this basis the activity coefficients of all electrolytes are unity at zero concentration. As the concentration is increased, the activity coefficients decrease at first below unity, pass through a minimum, and then increase again to values which may rise considerably above one. In Table 2 are listed the γ's for a number of electrolytes at 25° C, while in Fig. 1 is shown a plot for some of these as function of $\sqrt{m}$. Although at the higher concentrations the behavior of the γ's is highly individualized,

TABLE 2

MEAN ACTIVITY COEFFICIENTS, γ, OF ELECTROLYTES AT 25° C [1]

Molality m	HCl	LiCl	NaCl	KCl	HBr	NaOH	CaCl$_2$	H$_2$SO$_4$	Na$_2$SO$_4$	ZnSO$_4$	CuSO$_4$	La(NO$_3$)$_3$
0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.001	0.965	0.965	0.966	0.965	—	—	0.888	—	—	0.734	0.762	—
0.005	0.930	—	0.928	0.927	0.930	—	0.789	0.643	0.721	0.477	—	—
0.01	0.906	0.901	0.903	0.902	0.906	0.899	0.732	0.545	—	0.387	0.404	0.571
0.02	0.878	—	0.872	0.869	0.879	0.860	0.669	0.455	0.514	0.298	0.320	0.491
0.05	0.833	0.819	0.821	0.817	0.838	0.805	0.584	0.341	0.435	0.202	0.216	0.391
0.10	0.798	0.779	0.778	0.770	0.805	0.759	0.524	0.266	—	0.148	0.150	0.326
0.20	0.768	0.756	0.732	0.719	0.782	0.719	0.491	0.210	0.267	0.104	0.110	0.271
0.50	0.769	0.725	0.680	0.652	0.790	0.681	0.510	0.155	0.206	0.063	0.067	—
1.00	0.811	0.757	0.656	0.607	0.871	0.667	0.725	0.131	0.172	0.044	—	—
1.50	0.898	0.819	0.655	0.587	—	0.671	1.065	—	0.152	0.037	—	—
2.00	1.011	0.919	0.670	0.578	1.169	0.685	1.554	0.125	—	0.035	—	—
3.00	1.31	1.174	0.719	0.574	1.671	—	3.384	0.142	—	0.041	—	—
4.00	1.74	1.554	0.791	0.581	—	—	—	0.172	—	—	—	—

[1] Compiled from various sources, and particularly from MacInnes, "Principles of Electrochemistry," Reinhold Publishing Corporation, New York, 1939.

at the lower concentrations the curves seem to converge in a manner dependent on the type of electrolyte. This fact is of great significance in explaining theoretically the thermodynamic behavior of dilute solutions of strong electrolytes.
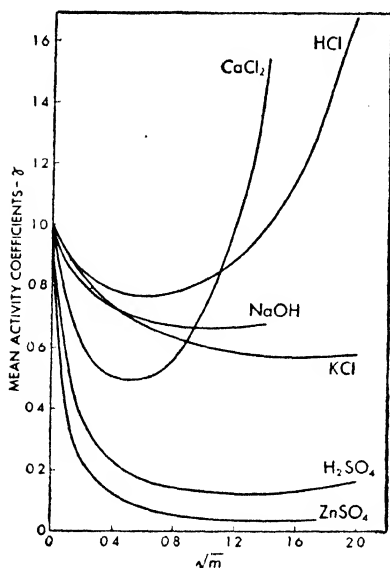
Fig. 1. Mean Activity Coefficients of Various Electrolytes at 25° C

The Debye-Hückel Theory of Activity Coefficients. The deviation of activity coefficients from unity is an indication that solutions of strong electrolytes are nonideal in their thermodynamic properties. Arrhenius tried to account for this deviation by postulating that the cause is partial dissociation. However, in light of present information strong electrolytes must be considered to be completely dissociated, at least in the more dilute solutions, and interionic attraction rather than partial dissociation must be responsible for the observed behavior.

A theory which attempts to account for the activity coefficients of electrolytes in terms of electrostatic attractions operating between ions in a solution is that of Debye and Hückel. These authors start by assuming that electrically charged particles in solution, like other charges, are subject to the Coulomb law of force, namely, that the forces of attraction or repulsion between charges vary directly as the product of the charges, q_1 and q_2, and inversely as the square of the distance r between them.

$$\text{Force} = \frac{1}{D}\left(\frac{q_1 q_2}{r^2}\right) \tag{22}$$

The proportionality constant D, called the *dielectric constant*, is determined by the medium in which the charges are immersed, in our case the solvent. As a result of these forces the distribution of ions throughout a solution is not random, but is such that any central positive ion is surrounded on an average by an atmosphere of other ions whose net charge is negative, while any central negative ion is surrounded by an atmosphere of net positive charge. Owing to the presence of an atmosphere about an ion there is a potential $\mathcal{E}_i$ established at its surface whose magnitude Debye and Hückel showed is

$$\mathcal{E}_i = \frac{-z_i e \kappa}{D(1 + \kappa a_i)} \tag{23}$$

In this equation z_i is the valence of the central ion, e the electronic charge, a_i the average ionic diameter, and κ is

$$\kappa = \sqrt{\frac{4\,\pi e^2 \Sigma n_i z_i^2}{DkT}} \tag{24}$$

where k is value of the gas constant per *molecule*, i.e., R/N, T the absolute temperature, and n_i is the number of any given kind of ions per *cubic centimeter* of solution. The summation Σ in equation (24) must be carried out over all the ions present in a given solution. κ has the dimensions of a reciprocal length, and may be looked upon as the reciprocal of the average thickness of the ionic atmosphere about an ion.

Now, the presence of a potential $\mathcal{E}_i$ at the surface of an ion due to the ionic atmosphere can be shown to give the solution an electrical free energy F_e,

$$F_e = \frac{-z_i^2 e^2 \kappa}{2\,D(1 + \kappa a_i)} \tag{25}$$

in excess of what the solution would have if the ionic atmosphere resulting from electrostatic attraction were not present. This excess free energy F_e is thermodynamically related to the activity coefficient of the ion f_i by

$$F_e = kT \ln f_i \tag{26}$$

and hence from equations (25) and (26) we obtain

$$\ln f_i = \frac{-z_i^2 e^2 \kappa}{2\,kTD(1 + \kappa a_i)} \tag{27}$$

On introducing into equation (27) the value of κ from equation (24), and the concentration of the ions C_i in gram ionic weights per liter from the identity,

$$C_i = \frac{n_i}{N} \times 1000$$

equation (27) becomes

$$\log_{10} f_i = \frac{-A z_i^2 \sqrt{\mu}}{1 + B a_i \sqrt{\mu}} \tag{28}$$

In this equation for any given temperature and any given solvent A and B are constants defined by

$$A = \frac{e^2}{2.303(DkT)^{3/2}} \sqrt{\frac{2\,\pi N}{1000}} \tag{29}$$

$$B = \sqrt{\frac{8\,\pi N e^2}{1000\,DkT}} \tag{30}$$

while μ is the ionic strength of the solution (see p. 221), i.e.,

$$\mu = \frac{1}{2} \Sigma C_i z_i^2 \qquad (31)$$

Equation (28) gives the *activity coefficient of an ion* as a function of the ionic strength of the solution. Since experimentally we evaluate not the ionic activity coefficient but the mean activity coefficient of the electrolyte f, equation (28) must be transformed with the aid of equation (7) to the logarithm of the mean activity coefficient. The result is

$$\log_{10} f = \frac{- A z_+ z_- \sqrt{\mu}}{1 + B a_i \sqrt{\mu}} \qquad (32)$$

The constants A and B for this equation are given for *water as a solvent* at various temperatures in Table 3. The mean ionic diameter a_i must be assumed to fit the data. However, *for very dilute solutions* $B a_i \sqrt{\mu}$ is small compared to unity and may be neglected. We obtain, then,

$$\log_{10} f = - A z_+ z_- \sqrt{\mu} \qquad (33)$$

which is the *limiting equation of Debye and Hückel for the activity coefficients of strong electrolytes.*

TABLE 3

DEBYE-HÜCKEL CONSTANTS A AND B FOR WATER AS SOLVENT

Temp. °C	A	B
0	0.4863	0.3243×10^8
18	0.4992	0.3272
25	0.5056	0.3286
38	0.5186	0.3314
100	0.568	0.342

Test of Debye-Hückel Theory. Because of certain simplifications necessitated by mathematical complexities of the derivation, the Debye-Hückel theory, as epitomized in equation (32), can be expected to be applicable only to *dilute solutions*. Further, equation (33) must be looked upon as a limiting law for the behavior of activity coefficients in *very dilute solutions*. According to equation (33) the activity coefficients of all strong electrolytes at high dilution should be determined, in a given solvent and at a given temperature, only by the ionic strength of the solution and the valence type ($z_+ z_-$) of the electrolyte, and not at all by the nature of the electrolyte. Further, a plot of $- \log f$ vs. $\sqrt{\mu}$ for all electrolytes should yield straight lines through the origin, but the slope of the lines should depend on $z_+ z_-$. For the 1–1 electrolytes this slope should be A, for 1–2 or 2–1 electrolytes $2A$, for 1–3 or 3–1 electrolytes

3 A, etc. That these requirements are met by experimented data in very dilute solutions may be seen from Fig. 2. In this figure are plotted $-\log f$ values of various complex cobalt ammines in aqueous solution as a function of the square root of the ionic strength. The points are the experimentally observed $-\log f$ values as obtained by Brönsted and LaMer [1] from solubility measurements, while the solid lines give the

$$\text{1-1} \quad [Co(NH_3)_4C_2O_4][Co(NH_3)_2(NO_2)_2C_2O_4]$$
$$\text{1-2} \quad [Co(NH_3)_4C_2O_4]_2[S_2O_6]$$
$$\text{3-1} \quad [Co(NH_3)_6][Co(NH_3)_2(NO_2)_2C_2O_4]_3$$
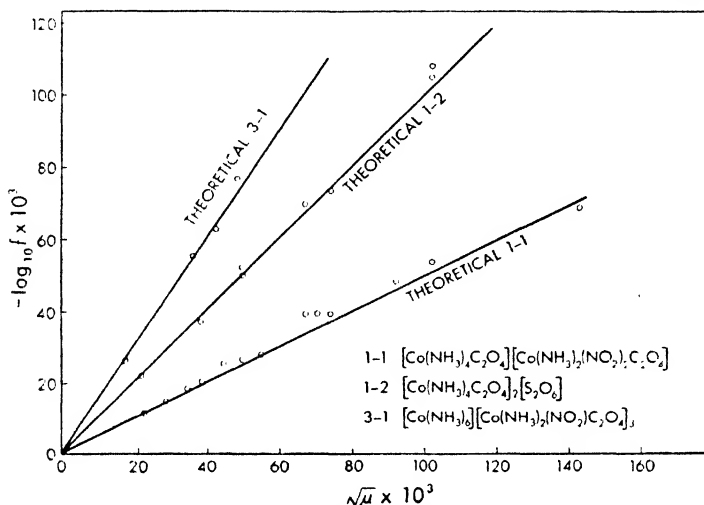
Fig. 2. Test of Debye-Hückel Limiting Law at 15° C

plots predicted by theory. Considering the difficulties involved in the determination of solubilities in such dilute solutions, the agreement between theory and experiment must be considered satisfactory. Similar tests have been carried out for other substances with the same result. We may conclude, therefore, that for strong electrolytes in extremely dilute solutions the limiting law of Debye and Hückel does represent the behavior of activity coefficients to $\sqrt{\mu} = 0.1$ or less for electrolytes of z_+z_- up to three. With valence types higher than $z_+z_- = 3$ the agreement usually does not extend to $\sqrt{\mu} = 0.1$.

For reproducing the course of activity coefficients with $\sqrt{\mu}$ at ionic strengths higher than $\sqrt{\mu} = 0.1$, the more complete equation (32) must be used. By an appropriate choice of a_i, which must be of the order of atomic diameters, i.e., several angstroms, the activity coefficients of 1–1 electrolytes can be reproduced to $C = 0.1$, and of 2–1 electrolytes to $C = 0.05$ or so. Hückel and others have pointed out that equation (32) can be extended even further by the introduction of another term linear in μ, but for our purposes this modification is not necessary, and will not be discussed.

[1] Brönsted and LaMer, J. Am. Chem. Soc., **46**, 555 (1924).

Although the Debye-Hückel theory was originally proposed for strong electrolytes, nothing in the theory would preclude its application to the *ionic portion of weak electrolytes*. This extension of the theory to weak electrolytes is a point of great importance, and will be illustrated in the course of the chapter.

The Ionization of Electrolytes. Arrhenius postulated that all electrolytes ionize in solution to a greater or lesser degree to yield ions in equilibrium with unionized molecules. If such equilibria do exist in all electrolytes, the law of mass action should be applicable to them, and it should be possible to calculate from the values of α as obtained from conductance or colligative measurements the ionization equilibrium constants, K.

Table 4 shows values of ionization constants calculated from conductance data for several strong (potassium chloride, sodium sulfate) and weak (acetic acid, ammonium hydroxide) electrolytes at 25° C. The results immediately reveal that whereas the K's for the weak electrolytes show very satisfactory constancy, the K's for the strong electrolytes do not meet the requirements of the law of mass action. The reason for the inapplicability of the law of mass action to the strong electrolytes is, of course, that for such electrolytes $\alpha = \Lambda/\Lambda_0$ measures not degree of ionization but interionic attraction in completely dissociated electrolytes. Therefore, there can be no question of applying the law of mass action to ionization of strong electrolytes, and we need concern ourselves only with the ionization equilibria of weak electrolytes, namely, the ionizations of weak acids and bases.

TABLE 4

IONIZATION CONSTANTS OF STRONG AND WEAK ELECTROLYTES AT 25° C

C	KCl	Na_2SO_4	$HC_2H_3O_2$	NH_4OH
0.0001			1.82×10^{-5}	1.79×10^{-5}
0.0005	0.0355	0.283×10^{-4}	1.82	1.81
0.001	0.0496	0.788	1.81	1.80
0.005	0.109	7.46	1.82	1.85
0.01	0.155	19.3	1.82	1.81
0.02	0.222	49.9	1.82	1.79
0.05	0.360	172.	1.82	1.80

Ionization Constants of Monobasic Acids. The ionization equilibrium of any weak monobasic acid, such as acetic acid, can be represented by the equation

$$HA = H^+ + A^-$$
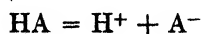
for which the *thermodynamic ionization constant* K_a is given by

$$K_a = \frac{a_{H^+} a_{A^-}}{a_{HA}} \tag{34}$$

Introducing the activity coefficients and concentrations for the activities, equation (34) becomes

$$K_a = \frac{(C_{H^+} f_{H^+})(C_{A^-} f_{A^-})}{(C_{HA} f_{HA})}$$

$$= \left(\frac{C_{H^+} C_A}{C_{HA}}\right)\left(\frac{f_{H^-} f_{A^-}}{f_{HA}}\right) \tag{35}$$

Now, for a total acid concentration C with degree of ionization α the concentration of HA will be $C_{HA} = C(1 - \alpha)$, while that of the ions will be $C_{H^+} = C_{A^-} = C\alpha$. On substitution of these into equation (35), we have

$$K_a = \frac{(C\alpha)(C\alpha)}{C(1 - \alpha)}\left(\frac{f_{H^-} f_{A^-}}{f_{HA}}\right)$$

$$= \left(\frac{C\alpha^2}{1 - \alpha}\right)\left(\frac{f_{H^-} f_A}{f_{HA}}\right)$$

$$= K_a' K_f \tag{36}$$

where

$$K_a' = \frac{C\alpha^2}{1 - \alpha} \tag{37}$$

is the *ionization constant of the acid in terms of concentrations*, while

$$K_f = \frac{f_{H^+} f_{A^-}}{f_{HA}} \tag{38}$$

is the ratio of the activity coefficients. For dilute solutions K_f will be close to unity, and for such solutions K_a will be essentially equal to K_a'. However, in more concentrated solutions a correction to K_a' for the deviation of the activity coefficients from unity will be in order.

Use of the above equations to calculate K_a' and K_a may be illustrated specifically with data on the ionization of acetic acid obtained from conductance measurements by MacInnes and Shedlovsky.[1] The first column in Table 5 lists the concentrations of the acid, while the second column gives the degree of ionization α of the acid as estimated from conductance, not from $\alpha = \Lambda/\Lambda_0$, but from the more exact equation (49) of the last chapter, i.e., $\alpha = \Lambda/\Lambda_e$. The concentration ionization constants K_a', calculated from C and α by means of equation (37), are given in the third column. It will be observed that although the K_a''s show a fairly satisfactory constancy, there is a consistent increase in K_a' with concen-

[1] MacInnes and Shedlovsky, J. Am. Chem. Soc., **54**, 1429 (1932).

tration. This drift is due to the neglect of K_f in equation (36). To take this factor into account we proceed as follows. On taking logarithms of this equation, we have

$$\log K_a = \log K_a' + \log K_f$$
$$= \log K_a' + \log \frac{f_{\mathrm{H^+}} f_{\mathrm{A^-}}}{f_{\mathrm{HA}}} \tag{39}$$

Since f_{HA} is the activity coefficient of an unionized molecule, it cannot be far from unity. Again, since $f_{\mathrm{H^+}} f_{\mathrm{A^-}} = f^2$ according to equation (7), then

$$\log f_{\mathrm{H^+}} f_{\mathrm{A^-}} = 2 \log f$$

and on introduction of the limiting law of Debye and Hückel, equation (33), we find for $2 \log f$ at 25° C with $z_+ = z_- = 1$

$$2 \log f = - 2 A z_+ z_- \sqrt{\mu}$$
$$= - 1.011 \sqrt{\mu} \tag{40}$$

The ionic strength of the solution is in this instance one-half the total *ionic concentration*, namely, $\mu = C\alpha$. Inserting this value of μ into equation (40), and substituting equation (40) into equation (39), we obtain finally, for $\log K_a$,

$$\log K_a = \log K_a' - 1.011 \sqrt{C\alpha} \tag{41}$$

The values of K_a calculated with equation (41) are shown in the last column of Table 5. It will be observed that within the range of validity of the limiting law of Debye and Hückel the values of K_a are now constant, and do not exhibit the drift shown by K_a'. Consequently 1.75×10^{-5} is the thermodynamic ionization constant of acetic acid at 25° C.

TABLE 5

THE IONIZATION CONSTANT OF ACETIC ACID AT 25° C

C	α	K_a'	K_a
0.00002801	0.5393	1.77×10^{-5}	1.75×10^{-5}
0.0001114	0.3277	1.78	1.75
0.0002184	0.2477	1.78	1.75
0.001028	0.1238	1.80	1.75
0.002414	0.08290	1.81	1.75
0.005912	0.05401	1.82	1.75
0.009842	0.04222	1.83	1.75
0.02000	0.02988	1.84	1.74
0.05000	0.01905	1.85	1.72
0.10000	0.01350	1.85	1.70

Although strictly speaking the ionization constants should always be expressed in activities, for many purposes no appreciable error is introduced by using K_a for K_a' or vice versa. Hence, we shall hereafter employ the practice of expressing the ionization constants in concentrations rather than activities for calculation purposes.

Ionization of Polybasic Acids. Acids possessing more than one ionizable hydrogen do not ionize in a single step, but in successive stages. Thus, sulfuric acid ionizes in two steps, namely,

$$H_2SO_4 = H^+ + HSO_4^-$$
$$HSO_4^- = H^+ + SO_4^{--}$$
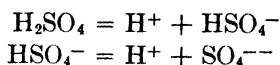
while phosphoric acid dissociates in three steps

$$H_3PO_4 = H^+ + H_2PO_4^-$$
$$H_2PO_4^- = H^+ + HPO_4^{--}$$
$$HPO_4^{--} = H^+ + PO_4^{---}$$
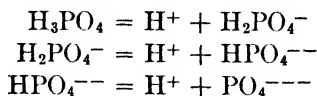
In any dibasic acid the primary ionization is more complete than the second, while in any tribasic acid the primary dissociation is greater than the second, and the second greater than the third. Each of the dissociation stages constitutes a true equilibrium, and for each of these there is an ionization constant. In sulfuric acid the first ionization stage is complete, i.e., H_2SO_4 is a strong acid. However, the secondary ionization of this acid, i.e., the ionization of HSO_4^-, is incomplete, and hence for this ionization stage we have the constant

$$K_a = \frac{a_{H^+} a_{SO_4^{--}}}{a_{HSO_4^-}}$$

In phosphoric acid all three stages involve partial ionization, and we have thus for this acid three ionization constants, K_{a_1}, K_{a_2}, K_{a_3}. At 25° C the magnitudes of the three constants are $K_{a_1} = 7.5 \times 10^{-3}$, $K_{a_2} = 6.2 \times 10^{-8}$, and $K_{a_3} = 1.7 \times 10^{-12}$, indicating in line with the above statement that at any given acid concentration the extent of dissociation of each stage decreases markedly from the primary to the tertiary. In fact, it can be calculated from the above ionization constants that, in a 0.0233 molar solution of phosphoric acid where the concentration of hydrogen ions is approximately 0.01, the concentrations of the various molecular and ionic species present are

$$C_{H_3PO_4} = 0.0133 \qquad C_{HPO_4^{--}} = 6 \times 10^{-8}$$
$$C_{H_2PO_4^-} = 0.01 \qquad C_{PO_4^{---}} = 1 \times 10^{-17}$$

From these figures it may be seen that even dilute solutions of pure phosphoric acid consist primarily of undissociated H_3PO_4, H^+, and $H_2PO_4^-$, and that in such solutions practically no PO_4^{---} ions are present.

Ionization Constants of Weak Bases. Any weak base BOH, such as ammonium hydroxide, ionizes according to

$$BOH = B^+ + OH^- \tag{42}$$

For this process the *thermodynamic ionization constant of the base K_b* is

$$K_b = \frac{a_{B^+}a_{OH^-}}{a_{BOH}} \tag{43}$$

while the *concentration ionization constant K_b'* is

$$K_b' = \frac{C_{B^+}C_{OH^-}}{C_{BOH}} \tag{44}$$

As in weak acids, the two constants K_b and K_b' are identical only in dilute solutions. Otherwise,

$$K_b = K_b'K_f \tag{45}$$

where K_f is again the activity coefficient ratio for the various species. We shall employ K_b' expressed in concentration units as a sufficiently close approximation to K_b.

Determination of Ionization Constants. Ionization constants of weak acids and bases can be determined from conductance measurements, and measurements of the hydrogen ion concentration of a given solution of the acid or base in presence of its salt. The conductance method has already been discussed in connection with the ionization constant of acetic acid. We turn our attention, then, to the estimation of ionization constants from hydrogen ion concentration measurements.

For this purpose consider specifically a solution of acetic acid of concentration C containing also sodium acetate at concentration C'. If in this solution the concentration of hydrogen ions is C_{H^+}, then the concentration of unionized acid is $C_{HA} = C - C_{H^+}$, while the concentration of acetate ions is that due to ionization plus that contributed by the completely ionized salt. Since on ionization of the acid equivalent quantities of hydrogen and acetate ions result, the acetate ion concentration resulting from ionization is C_{H^+}; and, this concentration plus C' from the salt gives $C_{A^-} = (C_{H^+} + C')$. In terms of these quantities, the ionization constant of the acid becomes

$$
\begin{aligned}
K_a' &= \frac{C_{H^+}C_{A^-}}{C_{HA}} \\
&= \frac{C_{H^+}(C_{H^+} + C')}{C - C_{H^+}}
\end{aligned} \tag{46}
$$

Therefore, as C and C' are known, measurement of C_{H^+} of the solution by some suitable means is all that is necessary to calculate K_a'.

Similar considerations apply to the determination of the ionization constants of bases. For the base ammonium hydroxide at concentration

C in presence of ammonium chloride at concentration C', the ionization constant K_b' takes the form

$$K_b' = \frac{C_{\mathrm{OH}^-}(C_{\mathrm{OH}^-} + C')}{C - C_{\mathrm{OH}^-}} \tag{47}$$

As will be shown presently, the concentrations of hydrogen and hydroxyl ions in water solution are related in a definite manner, so that as soon as C_{H^+} is measured, C_{OH^-} is also known. Hence, by measuring the *hydrogen ion* concentration in the solution of the base and its salt, C_{OH^-} is also measured, and K_b' follows from equation (47).

TABLE 6

IONIZATION CONSTANTS OF WEAK ACIDS AT 25° C

Acid	Formula	K_{a_1}	K_{a_2}	K_{a_3}
Arsenic	H_3AsO_4	5.0×10^{-3}	4.0×10^{-5}	6×10^{-10}
Boric	H_3BO_3	5.80×10^{-10}		
Carbonic	H_2CO_3	4.52×10^{-7}	5.59×10^{-11}	
Hydrocyanic	HCN	7.2×10^{-10}		
Iodic	HIO_3	1.9×10^{-1}		
Phosphoric	H_3PO_4	7.54×10^{-3}	6.23×10^{-8}	1.7×10^{-12}
Phosphorous	H_3PO_3	5×10^{-2}	2×10^{-5}	
Sulfuric	H_2SO_4	strong	1.20×10^{-3}	
Sulfurous	H_2SO_3	1.7×10^{-2}	5×10^{-6}	
Formic	$HCOOH$	1.77×10^{-4}		
Acetic	CH_3COOH	1.75×10^{-5}		
Propionic	C_2H_5COOH	1.34×10^{-5}		
Chloracetic	$CH_2ClCOOH$	1.38×10^{-3}		
Dichloracetic	$CHCl_2COOH$	5×10^{-2}		
Benzoic	C_6H_5COOH	6.6×10^{-5}		
Oxalic	$(COOH)_2$	3.8×10^{-2}	4.9×10^{-5}	
Phenol	C_6H_5OH	1.3×10^{-10}		

In Table 6 are listed the ionization constants of a number of weak acids at 25° C, while in Table 7 are given the ionization constants of some weak bases at the same temperature.

TABLE 7

IONIZATION CONSTANTS OF WEAK BASES AT 25° C

Base	Formula	K_b
Ammonia	NH_4OH	1.81×10^{-5}
Silver hydroxide	$AgOH$	1.1×10^{-4}
Methyl amine	$(CH_3)NH_2$	4.0×10^{-4}
Dimethyl amine	$(CH_3)_2NH$	7.4×10^{-4}
Trimethyl amine	$(CH_3)_3N$	7.4×10^{-5}
Ethyl amine	$(C_2H_5)NH_2$	5.6×10^{-4}
Aniline	$(C_6H_5)NH_2$	4.6×10^{-10}
Hydrazine	$NH_2 \cdot NH_2$	3×10^{-6}
Pyridine	$C_6H_5N \cdot$	2.3×10^{-9}
Urea	$CO(NH_2)_2$	1.5×10^{-14}

Calculations Involving Ionization Constants. When the ionization constant of an acid or base is known, it is readily possible to calculate the degree of ionization and the concentrations of the species present in solution under various given conditions. Two such calculations will be presented as examples of the procedure.

(a) Calculation of the degree of ionization of pure acids or bases.

Suppose that it is desired to calculate the degree of ionization and the concentrations of the various species in a solution of a monobasic acid at concentration C. If α is the degree of ionization, then according to equation (37)

$$K'_a = \frac{C\alpha^2}{1 - \alpha} \tag{37}$$

from which α follows as

$$\alpha = \frac{-K'_a + \sqrt{(K'_a)^2 + 4\,K'_a C}}{2\,C} \tag{48}$$

When K'_a is small, so will be α, and the denominator in equation (37) will be essentially unity. Under these conditions the equation for α reduces to

$$\alpha^2 = \frac{K'_a}{C}$$

$$\alpha = \sqrt{\frac{K'_a}{C.}} \tag{49}$$

Once α is known, C_{HA}, C_{H^+}, and C_{A^-} are readily obtained.

For the specific case of 0.01 molar propionic acid at 25° C, where $K'_a = 1.34 \times 10^{-5}$, we find for α, from equation (48),

$$\alpha = \frac{-1.34 \times 10^{-5} + \sqrt{(1.34 \times 10^{-5})^2 + 4(1.34 \times 10^{-5})0.01}}{2(0.01)}$$

$$= 0.0364$$

If equation (49) had been used, α would be

$$\alpha = \sqrt{\frac{K'_a}{C}} = \sqrt{\frac{1.34 \times 10^{-5}}{0.01}}$$

$$= 0.0366$$

With this value of α we find that $C_{H^+} = C_{A^-} = C\alpha = 0.01(0.0364) = 0.00036$, and that $C_{HA} = C(1 - \alpha) = 0.00964$.

In an exactly analogous manner may be calculated α and the concentrations in the ionization of bases.

(b) Degree of ionization in presence of a common ion.

When the solution of a weak acid or base contains as well a substance possessing an ion in common with the weak electrolyte, the degree of ionization of the latter is invariably repressed. The generality of this statement may be illustrated with the following example. Suppose that we have a C molar solution of propionic acid, and suppose, further, that this solution contains also sodium propionate at concentration C'. If α is again the degree of ionization, then $C_{HA} = C(1 - \alpha)$, and $C_{H^+} = C\alpha$. But the concentration of propionate ions is now that due to ionization, $C\alpha$, plus that contributed by the salt, C', or, $C_{A^-} = (C\alpha + C')$. Inserting these concentrations into the expression for K_a', we have

$$K_a' = \frac{C_{H^+}C_{A^-}}{C_{HA}}$$

$$= \frac{(C\alpha)(C\alpha + C')}{C(1 - \alpha)}$$

$$= \frac{\alpha(C\alpha + C')}{(1 - \alpha)} \tag{50}$$

and hence, $\quad \alpha = \dfrac{-(C' + K_a') + \sqrt{(C' + K_a')^2 + 4 K_a'C}}{2C} \tag{51}$

However, if we take $(1 - \alpha) = 1$, equation (51) becomes

$$\alpha = \frac{-C' + \sqrt{(C')^2 + 4 K_a'C}}{2C} \tag{52}$$

For a mixture of propionic acid again at $C = 0.01$ molar, and sodium propionate at $C' = 0.02$ molar, equation (52) yields for α

$$\alpha = \frac{-0.02 + \sqrt{(0.02)^2 + 4(1.34 \times 10^{-5})0.01}}{2(0.01)}$$

$$= 0.00070$$

Therefore, addition of 0.02 molar sodium propionate to 0.01 molar propionic acid reduces the ionization of the latter from $\alpha = 0.0366$, or 3.66 per cent, to $\alpha = 0.00070$, or 0.070 per cent.

With low values of K_a' no significant error is introduced when not only is α disregarded in the denominator of equation (50), but also $C\alpha$ as compared to C' in the numerator. When this is done, equation (50) becomes

$$K_a' = C'\alpha$$

and $\quad\quad\quad\quad\quad\quad \alpha = \dfrac{K_a'}{C'} \tag{53}$

For the case cited above, equation (53) would yield $\alpha = 0.00067$ as against 0.00070 by equation (52).

Another instance of common ion effect on the ionization of weak acids is the addition of a strong acid to the weak. In a solution containing 0.01 molar propionic acid and, say, 0.03 molar hydrochloric acid, we may take the hydrogen ion concentration as equal to that of the hydrochloric acid, namely, 0.03, and the concentrations of propionate ions and propionic acid as (0.01 α) and 0.01, respectively. Then,

$$K'_a = 1.34 \times 10^{-5} = \frac{(0.03)(0.01\ \alpha)}{0.01}$$

$$\alpha = 0.00045$$

and we see that the hydrogen ions of the strong acid, like the propionate ions of the salt, depress the ionization of the weak acid.

The Ion Product of Water. Conductance measurements and other evidence definitely point to the fact that water ionizes according to the equation

$$H_2O(l) = H^+ + OH^- \tag{54}$$
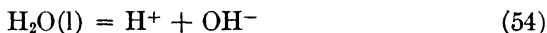
For this ionization, the equilibrium constant is

$$K = \frac{a_{H^+}a_{OH^-}}{a_{H_2O}} \tag{55}$$

However, since the ionization is at best very slight, the activity of the water in any aqueous solution will be constant, and may be included in K. Equation (55) becomes then

$$K_w = Ka_{H_2O} = a_{H^+}a_{OH^-} \tag{56}$$

The constant K_w is called the *ion product of water*. It indicates that in any aqueous solution both hydrogen and hydroxyl ions must be present, and, further, that at all times the product of the activities of the two ions must be a constant. In terms of concentrations this constant may be written as

$$K'_w = C_{H^+}C_{OH^-} \tag{57}$$

where K'_w differs from K_w by the product of the activity coefficients of the ions. In solutions of low ionic strength K'_w will not differ essentially from K_w.

The value of K'_w may be calculated from Kohlrausch and Heydweiller's data on the conductance of pure water given in the preceding chapter. These data yield for the degree of ionization of water at 25° C $\alpha = 1.9 \times 10^{-9}$. Now, as 1 liter of water at 25° C weighs 997.07 grams, and as the molecular weight of water is 18.016, the molar concentration C of

the water is $C = 997.07/18.016 = 55.34$ molar. From C and α, the concentrations of the ions follow as

$$C_{H^+} = C_{OH} = C\alpha$$
$$= 55.34 \times 1.9 \times 10^{-9}$$
$$= 1.05 \times 10^{-7}$$

and therefore K'_w at 25° C is

$$K'_w = C_{H^+}C_{OH^-}$$
$$= (1.05 \times 10^{-7})^2$$
$$= 1.10 \times 10^{-14}$$

In Table 8 are given the present best values for the thermodynamic ion product of water, K_w, at a number of temperatures. One point of interest in this table is that the ion product of water is 1×10^{-14} only at 25° C. At other temperatures the constant has different values, and consequently the concentration of hydrogen and hydroxyl ions in pure water at these temperatures will *not* be the same as at 25° C, namely, 1×10^{-7} gram ionic weights per liter.   .

TABLE 8

Ion Product of Water, K_w, at Various Temperatures

Temperature ° C	K_w
0	0.115×10^{-14}
10	0.293
25	1.008
40	2.916
60	9.614

As the ion product principle must be valid in any solution in which water is present, this constant may be employed to calculate the concentrations of hydrogen and hydroxyl ions present in such solutions. For instance, in 0.001 molar aqueous sodium hydroxide at 25° C $C_{OH^-} = 0.001$, and therefore the concentration of hydrogen ions in this solution must be

$$C_{H^+} = \frac{K'_w}{C_{OH^-}}$$
$$= \frac{1.01 \times 10^{-14}}{1 \times 10^{-3}}$$
$$= 1.01 \times 10^{-11}$$

Similarly the concentration of hydroxyl ions in any acid solution may be calculated.

Generalized Concept of Acids and Bases. As ordinarily employed, the term *acid* is taken to mean any substance that yields hydrogen ions

in solution, while a *base* is any substance that yields hydroxyl ions. Although these narrow definitions of acid and base may be satisfactory for some purposes in aqueous solutions, they are altogether insufficient to cover all observed phenomena either in water or in nonaqueous solvents.

A hydrogen ion has heretofore been considered to be a hydrogen atom with the electron removed, namely, the nucleus of the hydrogen atom or *proton*. However, evidence from various directions seems to point to the fact that a bare proton can have no existence as such in aqueous solutions. Brönsted [1] has shown that the free energy change of the reaction

$$H^+ + H_2O = H_3O^+$$
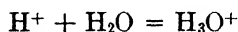
must be extremely large and negative, and hence the equilibrium constant of this reaction must also be very large. H^+ ions as such must be, then, practically nonexistent in aqueous media, and what we think of as the hydrogen ion is actually the hydrated proton, H_3O^+. The latter ion has variously been designated as the *hydronium, hydroxonium,* or *oxonium* ion. Likewise, certain studies of the glass electrode by Dole [2] and of the kinetics of reactions subject to acid catalysis have led to the conclusion that the proton is solvated.

These and a number of other considerations have led Brönsted [3] and Lowry [4] to a redefinition of the concept of acid and base. These authors define an *acid* as *any substance that can give, or donate, a proton to any other substance.* Again, they define a *base* as *any substance that can receive, or accept, a proton from an acid.* In other words, an acid is any substance, whether charged or uncharged, that can act as a *proton donor*, while any substance, whether charged or uncharged, that can act as a *proton acceptor* is a base. Inherent in these new definitions of acid and base is the significant fact that when an acid gives off a proton there must be a base to receive it; and, vice versa, no base can act as such unless there is an acid present to donate protons to it.

The differences between the newer *generalized concept of acids and bases* and the older, more restricted concept can best be brought out with several examples. According to both points of view acetic acid is an acid. However, whereas the older concept represents the ionization of this acid by the process

$$CH_3COOH = H^+ + CH_3COO^- \qquad (58a)$$

the newer concept represents the ionization as

$$CH_3COOH + H_2O = H_3O^+ + CH_3COO^- \qquad (58b)$$
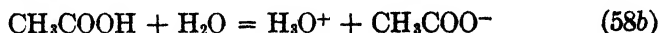
[1] Brönsted, Chem. Rev., **5**, 231 (1928).
[2] Dole, J. Am. Chem. Soc., **54**, 2120, 3095 (1932).
[3] Brönsted, *loc. cit.*; Rec. trav. chim., **42**, 718 (1923).
[4] Lowry, J. Chem. Soc., **123**, 848 (1923).

In the latter equation the acetic acid donates a proton to a water molecule, which acts as a *base*, to form a hydronium ion and an acetate ion. Furthermore, since H_3O^+ can donate a proton to the acetate ion to form acetic acid and water, H_3O^+ itself must be an acid, while CH_3COO^- *must be a base*. Consequently, any interaction of some one acid and some one base must always result in the formation of another acid and another base. Or, in general,

$$\text{Acid}_1 + \text{Base}_1 = \text{Acid}_2 + \text{Base}_2 \tag{59}$$
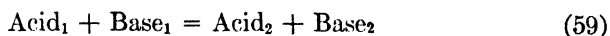
Base$_2$, which results from Acid$_1$, is said to be the base *conjugate* to Acid$_1$. Similarly Acid$_2$, which results from Base$_1$, is said to be *conjugate* to Base$_1$.

By extending this argument to other substances, we obtain the following formulations for the ionization in water of a number of *acids* in the generalized sense:

Acid$_1$		Base$_1$		Acid$_2$		Base$_2$
HCl	+	H_2O	=	H_3O^+	+	Cl^-
HCOOH	+	H_2O	=	H_3O^+	+	$HCOO^-$
HSO_4^-	+	H_2O	=	H_3O^+	+	SO_4^{--}
NH_4^+	+	H_2O	=	H_3O^+	+	NH_3
$C_6H_5NH_3^+$	+	H_2O	=	H_3O^+	+	$C_6H_5NH_2$
H_3O^+	+	H_2O	=	H_3O^+	+	H_2O
H_2O	+	H_2O	=	H_3O^+	+	OH^-
H_2SO_3	+	H_2O	=	H_3O^+	+	HSO_3^-
HSO_3^-	+	H_2O	=	H_3O^+	+	SO_3^{--}

From these examples it may be seen that besides the substances ordinarily considered to be acids, i.e., HCl, HSO_4^-, H_2SO_4, HCOOH, etc., H_3O^+, H_2O, NH_4^+, $C_6H_5NH_3^+$, and other proton donors are also acids. Again, bases are not only substances which possess hydroxyl ions, but also anions of acids, water, ammonia, aniline, HSO_3^-, and other proton acceptors. It will be observed that water may act as both acid or base, depending on the conditions and the reaction, i.e., water is *amphoteric*. Further, an ion like HSO_3^- is also amphoteric, for it may act as an acid or base, depending on circumstances of the reaction.

The strength of a given acid in the new theory is measured by its ability to donate protons to the solvent, and is expressible by the ionization constant of the acid. On the other hand, the strength of a given base K_B is defined as

$$K_B = \frac{1}{K_a} \tag{60}$$

where K_a is the ionization constant of the acid conjugate to the base. For *aqueous* solutions it can be shown that

$$K_B = \frac{K_w}{K_b} \tag{61}$$

K_b being the ionization constant of the base as defined by equation (43).

The generalized concept of acids and bases does not contradict the older views of these substances, but rather extends them. And, although the ideas involved may appear strange to one not accustomed to thinking in these terms, we shall see in a subsequent chapter that these ideas are very fruitful and permit the correlation and explanation of certain kinetic phenomena in aqueous solutions which would otherwise appear to be very puzzling.

Neutralization and Hydrolysis. The reaction between an acid and a base in aqueous solution always results in the formation of water and a salt. If this reaction were complete, exact neutralization of an acid by a base, or vice versa, would always give an exactly neutral solution; i.e., one containing no excess of hydrogen or hydroxyl ions. This is very nearly the case when a strong acid like hydrochloric is neutralized by a strong base like sodium hydroxide. However, when a weak acid like acetic is neutralized by sodium hydroxide, it is found that the final solution is not neutral but basic. Again, when a weak base like ammonium hydroxide is neutralized by hydrochloric acid, the final solution is acid. The basicity of the final solution in one case and acidity in the other are due to the tendency of the salt formed by neutralization to react with water and thereby reverse the neutralization. This tendency of salts when dissolved in water to react with the solvent and thereby reverse the neutralization process is called *hydrolysis*.

In considering the hydrolytic behavior of various salts, four cases may be distinguished, namely, (1) salts of strong acids and strong bases, (2) salts of weak acids and strong bases, (3) salts of strong acids and weak bases, and (4) salts of weak acids and weak bases. Each of these categories will be discussed now in turn.

Salts of Strong Acids and Strong Bases. Sodium chloride may be taken as an example of the hydrolytic behavior of a salt of a strong acid and a strong base. This salt exists in aqueous solution as sodium and chloride ions. If these two ions were to react with the water, the products would be hydrochloric acid and sodium hydroxide. However, since the latter two are also strong electrolytes, the products would again be sodium and chloride ions, and the hydrogen and hydroxyl ions would recombine to form water. In other words, the products of hydrolysis would be identical with the reactants, and there would be no change in the nature of the species in solution. We may say, therefore, that a salt of a strong acid and a strong base does not hydrolyze, and the solution of such a salt is essentially neutral.

Salts of Weak Acids and Strong Bases. When a salt of a weak acid and a strong base, such as sodium acetate, is dissolved in water, the cation of the base, i.e., the sodium ion, will not undergo hydrolysis for the reason given above. However, the anion of the weak acid, i.e., the

acetate ion, will react with water to form sufficient unionized acetic acid for the ionization constant of the acid to hold. The result is that the acetate ion undergoes the hydrolytic reaction

$$CH_3COO^- + H_2O = CH_3COOH + OH^-$$
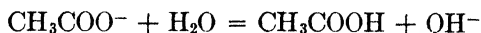
which leads also to the formation of hydroxyl ions, and the solution becomes alkaline.

In general, the hydrolysis of any salt BA of a weak acid HA and a strong base BOH is due to the hydrolysis of the anion of the acid. This hydrolysis of the anion can be represented by the equation

$$A^- + H_2O = HA + OH^- \tag{62}$$
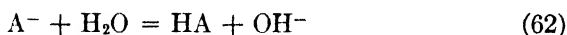
for which the equilibrium constant K_h, again including the activity of the water in the constant, is given by

$$K_h = \frac{a_{HA} a_{OH^-}}{a_{A^-}} \tag{63}$$

The constant K_h, called the *hydrolytic constant* of the ion A^-, determines the extent to which the ion A^- will react with water to form HA and OH^-. The magnitude of this constant, in turn, depends on the ionization constant of the acid HA, K_a, and the ion product of water K_w, as may be shown as follows. If the numerator and denominator of the right-hand side of equation (63) are multiplied by a_{H^+}, we obtain

$$K_h = \left(\frac{a_{HA}}{a_{A^-} a_{H^+}}\right)(a_{H^+} a_{OH^-})$$

But $\left(\dfrac{a_{HA}}{a_{A^-} a_{H^+}}\right) = \dfrac{1}{K_a}$, and $K_w = a_{H^+} a_{OH^-}$. Therefore,

$$K_h = \frac{K_w}{K_a} \tag{64}$$

and the hydrolytic constant may be calculated from the ion product of water and the ionization constant of the weak acid.

Salts of Strong Acids and Weak Bases. In salts of this class, of which ammonium chloride is an example, the anion of the strong acid will suffer no hydrolysis. But, the cation B^+ of the weak base BOH will undergo hydrolysis according to the reaction

$$B^+ + H_2O = BOH + H^+ \tag{65}$$
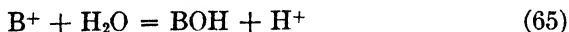
This time the hydrolysis yields molecules of the unionized base and hydrogen ions, and hence the solution of the salt in water is acid. The hydrolytic constant is now given by

$$K_h = \frac{a_{BOH} a_{H^+}}{a_{B^+}} \tag{66}$$

If the numerator and denominator of this expression are multiplied by a_{OH^-}, K_h becomes

$$K_h = \left(\frac{a_{BOH}}{a_{B^+}a_{OH^-}}\right)(a_{H^+}a_{OH^-})$$

But the first quantity in parentheses is $1/K_b$, where K_b is the ionization constant of the weak base BOH. Again, $K_w = a_{H^+}a_{OH^-}$. Therefore,

$$K_h = \frac{K_w}{K_b} \tag{67}$$

and the hydrolytic constant of a salt of a strong acid and a weak base can be calculated from the ion product of water and the ionization constant of the weak base BOH.

Salts of Weak Acids and Weak Bases. When the salt BA is the product of the interaction of a weak acid HA and a weak base BOH, such as ammonium acetate, both the cation and anion of the salt undergo hydrolysis. The reaction is

$$B^+ + A^- + H_2O = BOH + HA \tag{68}$$
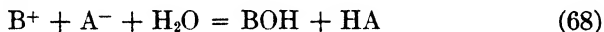
and whether the solution of the salt in water is acid or basic is determined by the relative strengths of the acid and base. In this instance the hydrolytic constant is defined by

$$K_h = \frac{a_{BOH}a_{HA}}{a_{B^+}a_{A^-}} \tag{69}$$

If now the numerator and denominator of equation (69) are multiplied by $a_{H^+}a_{OH^-}$, K_h becomes

$$K_h = \left(\frac{a_{BOH}}{a_{B^+}a_{OH^-}}\right)\left(\frac{a_{HA}}{a_{H^+}a_{A^-}}\right)(a_{H^+}a_{OH^-})$$

But, $\left(\dfrac{a_{BOH}}{a_{B^+}a_{OH^-}}\right) = \dfrac{1}{K_b}$, $\left(\dfrac{a_{HA}}{a_{H^+}a_{A^-}}\right) = \dfrac{1}{K_a}$, $(a_{H^+}a_{OH^-}) = K_w$. Therefore,

$$K_h = \frac{K_w}{K_b K_a} \tag{70}$$

This time the ionization constants of both the weak acid and the weak base must be available in order to determine the hydrolytic constant of the salt.

Calculations Involving Hydrolytic Constants. Hydrolytic constants of salts may be calculated from known K_a, K_b, and K_w values by means of equations (64), (67), or (70), or they may be determined experimentally. In the latter instance the usual practice is to dissolve a given amount of the salt in water, and measure the hydrogen ion con-

centration of the solution. Once this quantity and the original concentration of the salt are known, the hydrolytic constant may be calculated. In such calculations as well as in the converse process of calculating the degree of hydrolysis and the concentrations of the various species in the solution, it will be assumed that concentrations may be substituted for activities.

To illustrate the calculation of K_h from measurements of the hydrogen ion concentration of a salt in water, consider the following example. An 0.02 molar solution of sodium acetate in water at 25° C is found to have a hydrogen ion concentration of 3.0×10^{-9} gram ionic weights per liter. What is the hydrolytic constant of the salt? The hydrolytic reaction of this salt is given by

$$CH_3COO^- + H_2O = CH_3COOH + OH^-$$
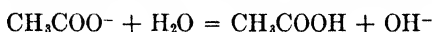
Since $C_{H^+} = 3.0 \times 10^{-9}$, and since $K_w = 1.01 \times 10^{-14} = C_{H^+}C_{OH^-}$, the concentration of hydroxyl ions must be

$$C_{OH^-} = \frac{K_w}{C_{H^+}} = \frac{1.01 \times 10^{-14}}{3.0 \times 10^{-9}}$$
$$= 3.37 \times 10^{-6}$$

This must also be the concentration of the acetic acid, as the latter is formed in quantity equivalent to the hydroxyl ions during the hydrolysis. Finally, the concentration of acetate ions at hydrolytic equilibrium must be the original concentration of the salt minus the amount reacted. Or,

$$C_{Ac^-} = 0.02 - 3.37 \times 10^{-6}$$
$$= 0.02$$

to a very near approximation. Therefore

$$K_h = \frac{C_{HA}C_{OH^-}}{C_{Ac^-}}$$
$$= \frac{(3.37 \times 10^{-6})(3.37 \times 10^{-6})}{0.02}$$
$$= 5.68 \times 10^{-10}$$

Calculated from $K_w = 1.01 \times 10^{-14}$ and $K_a = 1.75 \times 10^{-5}$ for acetic acid at 25° C, K_h is

$$K_h = \frac{K_w}{K_a}$$
$$= \frac{1.01 \times 10^{-14}}{1.75 \times 10^{-5}}$$
$$= 5.77 \times 10^{-10}$$

Once K_h is known, it may be employed to estimate the degree of hydrolysis of the salt under other conditions. Thus, suppose it is desired to know the degree of hydrolysis of sodium acetate in 0.01 molar solution at 25° C. If we let α be the degree of hydrolysis of the acetate ion, then $C_{Ac^-} = 0.01(1 - \alpha)$, and $C_{HAc} = C_{OH^-} = 0.01\,\alpha$. Consequently,

$$5.77 \times 10^{-10} = \frac{(0.01\,\alpha)(0.01\,\alpha)}{0.01(1 - \alpha)}$$

As K_h is small, so will be α, and we may write $1 - \alpha = 1$. Then,

$$\frac{(0.01)^2 \alpha^2}{0.01} = 5.77 \times 10^{-10}$$

$$\alpha = 2.40 \times 10^{-4}$$

i.e., the acetate ion is hydrolyzed to the extent of 0.024 per cent in 0.01 molar solution at 25° C.

pH and pOH. In passing from acid to alkaline solutions the concentration of hydrogen ions can vary within very wide limits. To permit a convenient means of expressing the concentration of hydrogen ions without involving negative exponents, Sørensen suggested the use of the *pH* (puissance d'hydrogen) *scale*. On this scale the pH of any solution is defined as

$$\text{pH} = -\log_{10} a_{H^+} \tag{71}$$

i.e., the hydrogen ion activity of a solution is equal to 10^{-pH}. Thus, for a solution of pH = 4, $a_{H^+} = 10^{-4}$, while for a solution of pH = 12, $a_{H^+} = 10^{-12}$. Similarly the activity of hydroxyl ions can be expressed on a *pOH scale* by the definition

$$\text{pOH} = -\log_{10} a_{OH^-} \tag{72}$$

Since in any aqueous solution $a_{H^+} a_{OH^-} = K_w$, we have

$$-\log_{10} a_{H^+} - \log_{10} a_{OH^-} = -\log_{10} K_w$$

and, therefore, in view of equations (71) and (72),

$$\text{pH} + \text{pOH} = -\log_{10} K_w \tag{73}$$

From this equation it may be seen that the sum of pH and pOH for any given aqueous solution must always be a constant equal to $-\log_{10} K_w$, and hence when the pH of a solution goes up the pOH must decrease, and vice versa. For equal concentrations of hydrogen and hydroxyl ions the solution will be neutral. For this condition pH must equal pOH, and therefore for neutrality

$$\text{pH} = \text{pOH} = \frac{-\log_{10} K_w}{2} \tag{74}$$

Any solution having a pH lower than $-1/2 \log_{10} K_w$ will thus be acid, while any solution having a pH higher than this value will be alkaline.

In the specific case of aqueous solutions *at 25° C*, $K_w = 1 \times 10^{-14}$, and $-\log_{10} K_w = 14$. For this temperature, then, pH + pOH = 14, and the neutral solution has a pH of seven. Any solution of pH lower than seven will be acid, while any solution of pH above seven will be alkaline.

Buffer Solutions. Solutions composed of an acid and one of its salts, or of a base and one of its salts, possess the ability to resist to a greater or lesser degree changes in pH of the solutions when some acid or base is added to them. Such solutions exhibiting the property of opposing a

change in their pH are called *buffer solutions,* or simply *buffers.* Particularly effective in this respect are mixtures of acid and salt or base and salt in which the acid or base is weak.

The buffering action of a solution of a weak acid HA in presence of one of its salts is explainable as follows. In any mixture of acid and salt the ionization equilibrium of the acid determines the hydrogen ion concentration of the solution. When some hydrogen ions are added to this solution in the form of some acid, the equilibrium is disturbed by the presence of an excess of hydrogen ions. To remove this excess and to reestablish the ionization equilibrium, hydrogen ions combine with anions of the salt to form molecules of unionized acid HA, and thereby the pH of the solution reverts to a value not far different from what it was originally. Again, when base is added to the mixture of acid and salt, the excess of hydroxyl ions disturbs the ionization equilibrium of the water. To reestablish this equilibrium hydrogen and hydroxyl ions combine to form water. As this reaction removes hydrogen ions, the ionization equilibrium of the acid is also disturbed, and to reestablish this equilibrium some of the acid HA ionizes to yield the requisite hydrogen ions. The result of these changes is that the excess hydroxyl ions are essentially neutralized, and the solution reverts to a pH close to its original value. These considerations apply only when the amount of acid or base added to a buffer solution is not so large as to change to any great extent the ratio of acid to salt.

In a like manner can be explained the buffering action of solutions of a weak base BOH and one of its salts. The particular pH at which a buffer solution is effective is determined by the ratio of acid or base to salt present and by the magnitude of the ionization constant of the acid or base. Thus, for a weak acid the hydrogen ion activity of the buffer is determined from the ionization constant by

$$a_{H^+} = K_a \frac{a_{HA}}{a_{A^-}} \qquad (75)$$

while for a weak base the hydroxyl ion activity is controlled by

$$a_{OH^-} = K_b \frac{a_{BOH}}{a_{B^+}} \qquad (76)$$

When the activity ratios of acid to salt or of base to salt are unity, these equations reduce to $a_{H^+} = K_a$ and $a_{OH^-} = K_b$; i.e., under these conditions the activities of H^+ and OH^- are given by the ionization constants only. However, when these activity ratios are not unity, the values of a_{H^+} and a_{OH^-} must be estimated from equations (75) and (76).

Buffer solutions are used whenever solutions of known pH are required, or whenever it is necessary to keep the pH of a given solution

constant. MacIlvaine [1] has given instructions for preparing buffer solutions of definite pH ranging from pH $= 3.4$ to pH $= 8.0$ in 0.2 unit steps. These are based on the use of disodium acid phosphate and citric acid solutions in various proportions. Another system of buffer mixtures ranging from pH $= 1$ to pH $= 10$ in 0.2 unit steps has been proposed by Clark and Lubs. Details for preparing these buffer mixtures may be found in Lange's Handbook [2] or in Clark. [3]

Acid-Base Indicators. Substances that undergo a change in color when the pH of a solution is changed are called *acid-base indicators*. Familiar to the student are such indicators as methyl orange, phenolphthalein, and litmus. These are merely a few of a large class of substances which possess the faculty of undergoing a transformation of color when acid is added to a basic solution or vice versa, and which possess characteristic colors within certain pH ranges. The particular pH at which an indicator undergoes a change in color is specific for each indicator. Again, the change in color does not take place abruptly, but occurs over several pH units. During this transformation of color the indicator takes on various hues intermediate between the characteristic acid and basic colors. This gradual change in color can be, and is, utilized for the determination of pH.

Acid-base indicators are weak organic acids or bases whose ionization depends on the pH of the solution. They differ from ordinary acids and bases in that the molecule of the unionized substance can exist in two tautomeric forms in equilibrium, in one of which the substance is a non-electrolyte, in the other a weak electrolyte capable of ionization. Thus the action of an acidic indicator can be represented by the scheme

$$\text{HIn} \;\rightleftharpoons\; \text{InH} \;\rightleftharpoons\; \text{In}^- + \text{H}^+ \tag{77}$$

<div align="center">Acid form   (Weak electrolyte)   Basic form
(Nonelectrolyte)   (Ionized)</div>
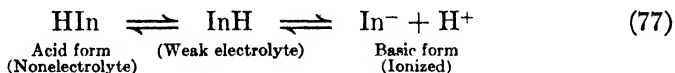
As all three forms of the indicator are in equilibrium, it is evident that addition of hydroxyl ions will favor the formation of the basic form, while addition of hydrogen ions will promote the formation of the acid form. Since these two forms of the indicator differ from each other in color, the latter will also change on addition of acid or base.

On applying the law of mass action to the equilibria shown in equation (77), we obtain for the equilibrium constant K_{In} of the overall process

$$K_{In} = \frac{[\text{In}^-][\text{H}^+]}{[\text{HIn}]} \tag{78}$$

[1] MacIlvaine, J. Biol. Chem., **49**, 183 (1921); see also "Lange's Handbook of Chemistry," fourth edition, Handbook Publishers, Inc., Sandusky, Ohio, 1941, p. 940.
[2] See "Lange's Handbook of Chemistry," p. 943.
[3] Clark, "The Determination of Hydrogen Ions," third edition, The Williams & Wilkins Co., Baltimore, 1928.

where the quantities in brackets designate the concentrations of the various species shown. This *indicator constant* determines the manner in which the concentrations of the acid and basic forms depend on each other and on the concentration of hydrogen ions in solution. With K_{In} available, it is readily possible to estimate the pH range over which a particular indicator will undergo change from acid to basic forms. As a rule the human eye is not very sensitive to color changes corresponding to $[HIn]/[In^-]$ much greater than 10 or less than 0.1. Therefore, the highest hydrogen ion concentration at which a change in color may be conveniently observed will be

$$[H^+] = K_{In} \frac{[HIn]}{[In^-]} = 10 \; K_{In}$$

while the lowest is

$$[H^+] = 0.1 \; K_{In}$$

This covers a pH interval of about two units over which color changes may be discerned for most indicators. In Table 9 are listed a number of indicators along with their indicator constants, pH ranges, and their acid and basic colors. From this list may be selected indicators sensitive over any desired pH interval.

TABLE 9

SOME ACID-BASE INDICATORS AND THEIR PROPERTIES

Indicator	K_{In}	pH Interval	Color Acid	Basic
Thymol blue	3.2×10^{-2}	1.2– 2.8	red	yellow
Methyl orange	2.0×10^{-4}	3.1– 4.4	red	yellow
Bromphenol blue	1.05×10^{-4}	3.0– 4.6	yellow	blue
Bromcresol green	2.14×10^{-5}	3.8– 5.4	yellow	blue
Chlorphenol red	1.05×10^{-6}	4.8– 6.4	yellow	red
Bromthymol blue	1.0×10^{-7}	6.0– 7.6	yellow	blue
Neutral red	1.41×10^{-7}	6.8– 8.0	red	yellow
Cresol red	5.0×10^{-9}	7.2– 8.8	yellow	red
Phenolphthalein	2.0×10^{-10}	8.3–10.0	colorless	red
Alizarin yellow	—	10.1–12.0	yellow	red
Indigo disulfonate	—	11.6–14.0	blue	yellow

Colorimetric Estimation of pH. Two general methods are available for the direct determination of the pH of a solution, namely, the *colorimetric* and the *potentiometric*. A discussion of the latter method will be deferred to the following chapter. The colorimetric method, on the other hand, entails the use of indicators, and will be described now.

The colorimetric method for determination of pH involves the comparison of the color developed by addition of a suitable indicator to the

solution whose pH is sought with the color of a series of standards of known pH prepared in the same manner as the test solution. The procedure involved can best be explained with a specific example. Suppose the pH of a particular solution is to be determined, and suppose, further, that it is known that the pH of the solution is somewhere near 7. Reference to Table 9 shows that a suitable indicator for this solution is bromthymol blue, whose color variation from yellow to blue may be followed over the pH interval 6.0 to 7.6. To obtain the standards containing this indicator, a series of buffers are prepared ranging in pH from 6.0 to 7.6 in, say, 0.2 pH units. To 10 ml of each of these buffers are added 10 drops of a previously prepared solution of the indicator. We have thus nine standard solutions, each of definite pH, and colors ranging from yellow for pH = 6 to blue for pH = 7.6. Next 10 ml of the test solution are taken, again 10 drops of the indicator are added, and the test solution and standards are compared to see which of the standard solutions matches that of the test solution in color. The pH of the unknown solution is then the pH of the standard whose color is identical with that of the test solution.

The various schemes proposed for the colorimetric estimation of pH are based on this principle. They differ from each other only in the manner in which the color comparisons are made. The simplest comparator is a block of wood drilled vertically to accommodate the test tubes of standard and test solutions, and horizontally for viewing the test tubes side by side. A better arrangement is the roulette comparator, where the standards are mounted in a rotating wheel so that various solutions can be brought into view for comparison with the unknown with a minimum of manipulation. More elaborate color comparators are the various colorimeters, such as the wedge, Duboscq, or Klett, where optical systems are utilized to bring the two solutions into view under controlled lighting conditions for a closer and more precise color comparison. All these methods depend on the human eye for the color match. However, in photoelectric colorimeters coming more and more into use, photoelectric cell arrangements permit a color comparison to be made without visual observation.

The accuracy with which the pH of a solution can be determined colorimetrically depends greatly on the nature of the solution and the comparator used. With an ordinary block comparator the pH can be estimated only to 0.1 to 0.2 pH unit. With optical and particularly with photoelectric colorimeters the accuracy can be improved. However, certain errors inherent in the colorimetric method cannot be obviated by refinements in observation. It has been found that high concentrations of salts or proteins in a solution lead to erroneous values for the pH. These salt and protein errors arise from the ionic strength effects these

substances exert on the indicator equilibrium. Unless corrected for by appropriate calibration, the measured pH's may be off anywhere from several tenths up to as much as 1 pH unit from the true values.

Still another difficulty arises when the test solution has a color of its own. Under such conditions special techniques must be employed to compensate for the color of the solution.

Other Homogeneous Ionic Equilibria. Ionic equilibria in homogeneous systems are not limited to ionizations and hydrolyses. Many other types of equilibrium are possible, which may involve only ions, as in

$$2 \, Fe^{+++} + Sn^{++} = 2 \, Fe^{++} + Sn^{++++}$$
$$Ag^+ + 2 \, CN^- = Ag(CN)_2^-$$
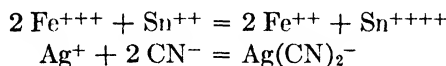
or ions and unionized substances, as in

$$I^- + I_2 = I_3^-$$
$$CH_3CH_2NO_2 + OH^- = CH_3CH = NO_2^- + H_2O$$
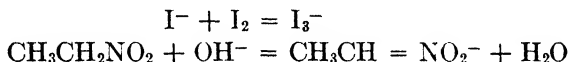
All such equilibria are handled in exactly the same manner as the other equilibria discussed. Through appropriate means the equilibrium constants of such ionic reactions can be evaluated, and these in turn may be used for calculating the activities or concentrations of the substances involved under various specified conditions.

Heterogeneous Ionic Equilibria. So far only ionic equilibria in homogeneous systems were mentioned. However, ions may be involved in heterogeneous equilibria, in which two or more phases are present, as well. As examples of such equilibria may be cited the reactions:

$$NH_3(g) + H_2O(l) = NH_4^+ + OH^- \qquad (a)$$
$$Sn(s) + 2 \, H^+ = Sn^{++} + H_2(g) \qquad (b)$$
$$AgBr(s) + I^- = AgI(s) + Br^- \qquad (c)$$
$$Sn(s) + Pb^{++} = Pb(s) + Sn^{++} \qquad (d)$$
$$BaSO_4(s) = Ba^{++} + SO_4^{--} \qquad (e)$$
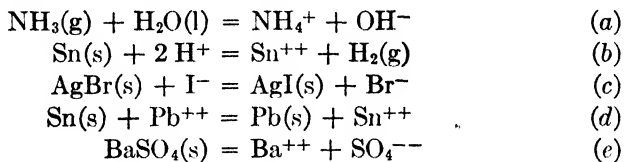
The first of these reactions represents the solubility of a gas in water to form a weakly ionized electrolyte, the second the solubility of a metal in an acid, the third and fourth a metathesis, and the last the solubility of a salt. Further, the equilibria represented in these reactions involve two phases, as in the first and last, or three phases, as in the second, third, and fourth reactions.

The formulation of the equilibrium constants of heterogeneous reactions involving ions introduces nothing new. All that need be remembered is that the activity of pure solid and liquid phases is unity, while pressure will be approximately equal to the activity of gas phases.

With this in mind, the expressions for the equilibrium constants of, say, reactions (a) and (c) become:

$$K_{(a)} = \frac{a_{NH_4^+} a_{OH^-}}{P_{NH_3}}$$

$$K_{(c)} = \frac{a_{Br^-}}{a_{I^-}}$$

As an example for discussion may be taken the work of Noyes and Toabe [1] on the equilibrium between tin and lead ions given in (d). These authors studied this equilibrium by agitating solutions of lead perchlorate, tin perchlorate, or mixtures of the two, with solid lead and tin, and then determined the concentrations of the stannous and plumbous ions at equilibrium. In Table 10 are given some of their results for 25° C. Columns 1 and 2 give the concentrations of the two ions at equilibrium, column 3 the equilibrium constant of the reaction expressed in molarities. The fair constancy of the K's confirms the correctness of the equation for the equilibrium constants and indicates also that in this instance the constant expressed in concentrations represents a fairly close approximation to the thermodynamic constant expressed in activities. The reason is that both ions are divalent, and hence the ratio of their activity coefficients, as it appears in the equation, is not far from unity.

TABLE 10

EQUILIBRIUM AT 25° C IN THE REACTION

$$Sn(s) + Pb^{++} = Pb(s) + Sn^{++}$$

Equilibrium Concentration

(moles/liter)		$K = \dfrac{C_{Sn^{++}}}{C_{Pb^{++}}}$
$C_{Pb^{++}}$	$C_{Sn^{++}}$	
0.0132	0.0413	3.14
0.0148	0.0457	3.08
0.0233	0.0704	3.02
0.0235	0.0692	2.95
0.0275	0.0821	2.98

The Solubility Product. A particularly important type of heterogeneous ionic equilibrium is involved in the solubility of difficultly soluble salts in water. When a difficultly soluble salt, such as barium sulfate, is agitated with water until the solution is *saturated*, the equilibrium established between the solid phase and the completely ionized salt in solution is given by

$$BaSO_4(s) = Ba^{++} + SO_4^{--} \tag{79}$$
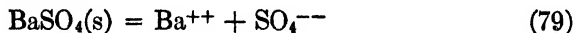
For this process the equilibrium constant is

$$K_s = a_{Ba^{++}} a_{SO_4^{--}} \tag{80}$$

[1] Noyes and Toabe, J. Am. Chem. Soc., **39**, 1537 (1917).

i.e., *in any solution saturated with barium sulfate the product of the activities of barium and sulfate ions is a constant equal to K_s.* The constant K_s is called the *solubility product constant*, or simply the *solubility product*, of the salt.

Every difficultly soluble salt has a solubility product constant of its own. For salts yielding only two ions, such as barium sulfate, silver chloride, lead sulfate, mercuric sulfide, the solubility product is merely the product of the activities of the two ions. However, for salts yielding more than two ions the solubility product expression is a little more complicated. In general, the solubility equilibrium of any salt A_xB_y yielding x positive ions of A and y negative ions of B is given by

$$A_xB_y(s) = xA + yB \qquad (81)$$

and hence the most general expression for the solubility product is

$$K_s = a_A^x a_B^y \qquad (82)$$

Thus, for a salt such as silver carbonate, for which $x = 2$ and $y = 1$, $K_s = a_{Ag^+}^2 a_{CO_3^{--}}$. Again, for calcium fluoride $x = 1$, $y = 2$, and $K_s = a_{Ca^{++}} a_{F^-}^2$. In these instances the expressions given will be constant for saturated solutions of the salts, and not merely the ionic products as in salts yielding two ions.

The solubility products are readily evaluated from the saturation solubilities of salts in pure water. These solubilities may be obtained by direct analysis of the saturated solutions, by conductance measurements of the saturated solutions as described in the last chapter, or by electromotive force measurements. The calculations involved may be illustrated with the following example. At $25°$ C the solubility of silver carbonate in water is 1.16×10^{-4} moles per liter. Hence the concentration of carbonate ions in the saturated solution is $C_{CO_3^{--}} = 1.16 \times 10^{-4}$, while that of the silver ions is twice that of the carbonate, or $C_{Ag^+} = 2.32 \times 10^{-4}$ gram ionic weights per liter. For silver carbonate K_s is given by

$$\begin{aligned} K_s &= a_{Ag^+}^2 a_{CO_3^{--}} \\ &= (C_{Ag^+}^2 C_{CO_3^{--}})(f_{Ag^+}^2 f_{CO_3^{--}}) \end{aligned}$$

If we assume now that the activity coefficients are unity in the very dilute solution involved, then K_s reduces to K_s', the stoichiometric solubility product. Therefore,

$$\begin{aligned} K_s = K_s' &= C_{Ag^+}^2 C_{CO_3^{--}} \\ &= (2.32 \times 10^{-4})^2(1.16 \times 10^{-4}) \\ &= 6.2 \times 10^{-12} \end{aligned}$$

In Table 11 are listed the solubility products at 25° C for a number of salts as well as for several difficultly soluble hydroxides to which the solubility product principle is applicable. From these constants the solubility of these substances in water can readily be obtained by reversing the calculations for the solubility product given above.

TABLE 11

SOLUBILITY PRODUCTS FOR VARIOUS SUBSTANCES AT 25° C

Substance	K_s	Substance	K_s
$Al(OH)_3$	3.7×10^{-15}	HgI_2	3.2×10^{-29}
$BaSO_4$	1.08×10^{-10}	$AgBr$	7.7×10^{-13}
$CaCO_3$	8.7×10^{-9}	$AgCl$	1.56×10^{-10}
CuS (18°)	8.5×10^{-45}	AgI	1.5×10^{-16}
$Fe(OH)_3$ (18°)	1.1×10^{-36}	Ag_2CO_3	6.15×10^{-12}
$Fe(OH)_2$ (18°)	1.64×10^{-14}	Ag_2CrO_4	9×10^{-12}
PbI_2	1.39×10^{-8}	Ag_2S	1.6×10^{-49}
$Mg(OH)_2$ (18°)	1.2×10^{-11}	$SrCO_3$	1.6×10^{-9}
$HgBr_2$	8×10^{-20}	$TlCl$	2.0×10^{-4}

Effect of Common Ions on Solubility. For any solution saturated with, say, thallous chloride, the solubility product principle demands that $a_{Tl^+}a_{Cl^-} = K_s$, or approximately $C_{Tl^+}C_{Cl^-} = K_s'$, *no matter what the source of the two ions may be.* Suppose now that to a saturated solution of this salt in water be added some potassium chloride. Addition of the potassium chloride increases the concentration of chloride ions, $C_{Tl^+}C_{Cl^-}$ becomes greater than K_s', and hence some thallous chloride must be precipitated from solution to reestablish the saturation equilibrium. The net result of the increase in the chloride ion concentration is, therefore, a decrease in the solubility of thallous chloride. Similarly, if a thallous salt, such as thallium nitrate, is added to a saturated solution of thallous chloride, the solubility of the latter must again decrease to preserve the constancy of the solubility product. This effect of common ions is operative in all solutions saturated with difficultly soluble salts. Consequently, we may state as a general rule that *addition of substances possessing ions in common with a dissolving salt causes a decrease in the solubility of the salt.* This rule is valid as long as the dissolving salt does not form a complex of some kind with the added substance. When complex formation does occur, the solubility of the salt may actually increase in presence of common ions. Thus, the solubility of $HgBr_2$ in water is less than in solutions containing bromides due to formation of $HgBr_3^-$. Since formation of the complex removes mercuric and bromide ions, more $HgBr_2$ must dissolve to satisfy the requirements of the solubility product principle.

The solubility of difficultly soluble salts in presence of common ions can be calculated from the solubility product. For this purpose we shall assume that $K_s = K_s'$, and express the constant in concentrations. Taking again thallous chloride, we have for this salt $K_s = 2.0 \times 10^{-4}$ according to Table 11. To obtain from this constant the solubility of thallous chloride in presence of added chloride ions, we proceed as follows. If C is the concentration of chloride ions due to added soluble chloride, such as sodium chloride, potassium chloride, or hydrochloric acid, and S the solubility of thallous chloride in presence of added substance, then the concentration of thallous ions at saturation will be S, that of chloride $(C + S)$, and the expression for the solubility product becomes

$$K_s' = C_{Tl^+}C_{Cl^-}$$
$$= S(C + S) \qquad (83)$$

With K_s' known, S may be calculated from equation (83) for any given value of C. Again, if the added salt contains thallous instead of chloride ions at a concentration C', the concentration of thallous ions will be $(C' + S)$, that of chloride ions S, and the expression for K_s' will become

$$K_s' = (C' + S)S \qquad (84)$$

Equations (83) and (84) indicate that for the same value of C and C' the solubility of thallous chloride should be the same; i.e., the common ion effect for a given added concentration of chloride or thallous ions should be the same for both ions and should be dependent only on the concentration.

The lowest curve in Fig. 3 shows the solubility of thallous chloride

Fig. 3. Effect of Salts on Solubility of TlCl at 25° C

predicted by equations (83) and (84) for various values of C or C'. The solubility of this salt actually observed in presence of added potassium chloride and thallium nitrate is shown in the next two curves. These data indicate that, although in general calculated and observed solubilities are parallel, the observed solubilities are higher than predicted by the approximate equations (83) and (84). Further, the common ion effect is not only dependent on the concentration, but depends also on the nature of the added substance. The reason for this behavior lies in what is known as the *salt effect*, to which we now turn.

Salt Effects and Solubility. From the formulation of the solubility product in terms of concentrations it is to be expected that only substances possessing an ion in common with the dissolving salt should affect the solubility, and that electrolytes not possessing a common ion should be without influence. Actually it is found that such electrolytes exercise a very pronounced effect, causing an *increase in solubility* with increase in concentration of added substance. This may be seen from the upper three curves in Fig. 3, where the effect of various concentrations of added sodium acetate, zinc sulfate, and potassium sulfate on the solubility of thallous chloride is depicted.

Although inexplicable from the concentration approach to the solubility product, this marked increase in solubility with electrolyte concentration is just what is to be expected from the more exact activity formulation of the solubility product. According to equation (82), the thermodynamic solubility product of thallous chloride should be

$$K_s = a_{Tl^+} a_{Cl^-} \tag{85}$$

Introducing concentrations and activity coefficients for the activities, equation (85) becomes

$$K_s = (C_{Tl^+} C_{Cl^-})(f_{Tl^+} f_{Cl^-}) \tag{86}$$
$$= K_s' f^2$$

where f is the mean activity coefficient of the salt, and K_s' the concentration solubility product. From equation (86) it follows that, since K_s must be a true constant, K_s' cannot be constant and equal to K_s unless $f = 1$. However, f is not unity except at zero ionic strength, but varies with ionic strength as the latter is changed. Consequently, K_s' must also be a function of the ionic strength and must vary with it according to the dependence on μ of f^2 in the expression

$$K'_s = \frac{K_s}{f^2} \tag{87}$$

Now, at not too excessively high ionic strengths the activity coefficients are less than one, and consequently at these ionic strengths K_s' must be greater than K_s. Since the solubility is equal to $\sqrt{K_s'}$, the solubility in presence of added salt must also be greater. We see, therefore, that whereas the thermodynamic solubility product K_s remains constant, theory predicts that the concentration solubility product K_s' should increase with increase in ionic strength, and so should the solubility, as is the case. This argument accounts also for the fact that the observed solubilities in presence of common ions are greater than predicted when no allowance is made for activity coefficients. Since the common ion and salt effects operate against each other, the former to decrease and the

latter to increase the solubility, and as the common ion effect is the more pronounced, the result of addition of common ions to a dissolving salt is to lower the solubility, but not quite to the extent which would be anticipated on the basis of no salt effect.

From these considerations it is evident that the solubility product principle is strictly valid only for K_s expressed in activities, and that *exact* calculations can be made only when activity coefficients are included. When activity coefficients at various ionic strengths are available for a dissolving salt, exact calculations of solubility introduce no difficulty. Thus, in presence of 0.1 molar potassium chloride the activity coefficient of thallous chloride is 0.715. Using $K_s = 2.0 \times 10^{-4}$ given in Table 11 and this activity coefficient, K_s' follows from equation (87) as

$$K_s' = \frac{K_s}{f^2}$$
$$= \frac{2.0 \times 10^{-4}}{(0.715)^2}$$
$$= 3.9 \times 10^{-4}$$

Now, in presence of added chloride the solubility of thallous chloride is given by equation (83). For $C = 0.10$ and $K_s' = 3.9 \times 10^{-4}$, this equation gives $S = 0.0039$ mole per liter as against the observed 0.00396. Calculated without activity coefficients a solubility of 0.0020 would have been obtained.

Activity Coefficients from Solubility Measurements. Solubility measurements in presence of added electrolytes may be employed to evaluate the activity coefficients of the *dissolving salt* at various ionic strengths. The procedure involved may again be illustrated with thallous chloride. If we represent the solubility of this salt in pure water by S_0, in presence of added electrolyte with no common ion by S, and let f_0 and f be the mean activity coefficients of the salt in the two solutions, then, in view of the constancy of the thermodynamic solubility product, we have the relation

$$K_s = S_0^2 f_0^2 = S^2 f^2 \tag{88a}$$

and therefore, $\qquad\qquad Sf = S_0 f_0 \tag{88b}$

Taking logarithms, we have

$$\log S + \log f = \log S_0 f_0$$
$$\log S = \log S_0 f_0 - \log f \tag{89}$$

According to this expression f can be evaluated from the solubilities as soon as the product $S_0 f_0$ is available. To obtain this product $\log S$ is plotted against the square root of the total ionic strength of the solution, i.e., of dissolved salt and added electrolyte, and the plot extrapolated to $\sqrt{\mu} = 0$. As at zero ionic strength $f = 1$, the intercept on the $\log S$ axis

gives immediately log $S_0 f_0$. Once this quantity is known, subtraction of log $S_0 f_0$ from log S yields $- \log f$, and hence f, at various ionic strengths. Or, on evaluating $S_0 f_0$, f follows from equation (88b) by dividing $S_0 f_0$ by S.

This procedure for evaluating f from solubility measurements may be illustrated with the data given in Table 12 on the solubility of thallous chloride at 25° C in water and in presence of various concentrations of potassium nitrate. Column 1 gives the concentration C of added salt, column 2 the observed solubility S of thallous chloride in moles per liter, and column 3, μ, which in this case is equal to $(C + S)$. From a plot of log S against $\sqrt{\mu}$ it is found that $S_0 f_0 = 0.01422$. Dividing now $S_0 f_0$ by S at the various ionic strengths, the activity coefficients in column 4 result.

TABLE 12

SOLUBILITY OF TlCl IN PRESENCE OF KNO₃ AT 25° C

Conc. KNO₃ Added C	Solubility TlCl S	$\mu = (C + S)$	Mean Activity Coefficient of TlCl f	$K_s = S^2 f^2$
0	0.01607	0.01607	0.885	2.02×10^{-4}
0.02	0.01716	0.03716	0.829	2.02×10^{-4}
0.05	0.01826	0.06826	0.779	2.02×10^{-4}
0.10	0.01961	0.11961	0.725	2.02×10^{-4}
0.30	0.02312	0.32313	0.615	2.02×10^{-4}
1.00	0.03072	1.03072	0.463	2.02×10^{-4}

For a final check on the calculations the last column gives the thermodynamic solubility product of thallous chloride, which according to equation (88a) is $K_s = S^2 f^2$. As it should be, K_s is constant throughout.

Ionic Equilibria and Temperature. Equilibrium constants of ionic reactions, like all other equilibrium constants, vary with temperature. As before, the variation of these equilibrium constants with temperature is given by

$$\frac{d \ln K}{dT} = \frac{\Delta H}{RT^2} \tag{90}$$

and this expression may be employed to evaluate the heats of various ionic reactions, or to estimate K at one temperature from that at another when ΔH is known.

REFERENCES FOR FURTHER READING

1. For a discussion of activity coefficients of electrolytes and of Debye-Hückel theory, see references listed at end of Chapter VII, as well as
   (a) Lewis and Randall, "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill Book Company, Inc., New York, 1923.

(b) F. H. MacDougall, "Thermodynamics and Chemistry," third edition, John Wiley & Sons, Inc., New York, 1939.

(c) L. E. Steiner, "Introduction to Chemical Thermodynamics," McGraw-Hill Book Company, Inc., New York, 1941.

2. H. T. S. Britton, "Hydrogen Ions," second edition, D. Van Nostrand Company, Inc., New York, 1932.

3. W. M. Clark, "The Determination of Hydrogen Ions," third edition, The Williams & Wilkins Company, Baltimore, 1928.

4. Creighton and Koehler, "Principles and Applications of Electrochemistry," John Wiley & Sons, Inc., New York, 1935, Vol. I.

5. K. Jellinek, "Lehrbuch der physikalischen Chemie," Ferdinand Enke, Stuttgart, 1930, Vol. IV.

6. I. M. Kolthoff, "Colorimetric and Potentiometric Determination of pH," John Wiley & Sons, Inc., New York, 1931.

7. Kolthoff and Furman, "Indicators," John Wiley & Sons, Inc., New York, 1926.

8. Kolthoff and Laitinen, "pH and Electrometric Titrations," second edition, John Wiley & Sons, Inc., New York, 1941.

PROBLEMS

Note: In problems where insufficient data are available for determining activities of ions of strong electrolytes, assume ionic activity equal to ionic concentration.

1. For each of the following solutions evaluate the mean molality, the mean ionic activity, and the activity of the salt:

	Molality	Mean Activity Coefficient
$K_3Fe(CN)_6$	0.01	0.571
$CdCl_2$	0.1	0.219
H_2SO_4	0.05	0.397

2. The density of a 1.19 molar solution of $CaCl_2$ in water is 1.10 g/millimeter at 20° C. What is the molality of the solution and the ratio of f to γ?

Ans. 1.228 molal; $f/\gamma = 1.029$

3. Using the Debye-Hückel limiting law, calculate the mean ionic activity coefficient of a 0.001 molal solution of $K_3Fe(CN)_6$, and compare with the observed value of 0.808.

4. A solution is 0.002 molal in $CoCl_2$ and 0.002 molal in $ZnSO_4$. Calculate the activity coefficient of Zn^{++} ions in the solution using the Debye-Hückel limiting law.

Ans. 0.587

5. The equivalent conductance of a 0.0140 N solution of chloracetic acid is 109.0 at 25° C. If Λ_0 is 389.5, what is the ionization constant of chloracetic acid?

Ans. 1.52×10^{-3}

6. From the following data calculate and compare the apparent ionization "constants" of HCl at the various concentrations:

Concentration	Λ	Concentration	Λ
0.000000	426.16	0.001877	419.76
0.000319	423.55	0.002994	418.10
0.000754	421.78		

7. Assuming that the conductance measurements give the true degree of dissociation in problem 5, use the Debye-Hückel limiting law to calculate the thermodynamic dissociation constant of chloracetic acid.

8. Find at 25° C the ratio of the degree of dissociation of a 0.1 molar aqueous solution of methylamine to that of a 0.001 molar solution.

9. Calculate the degree of dissociation and the hydrogen ion concentration in a 0.05 molar chloracetic acid solution at 25° C.

10. Calculate the H^+, $H_2PO_4^-$, HPO_4^{--}, and PO_4^{---} ion concentrations in a 0.1 molar H_3PO_4 solution at 25° C.

11. Compare for 25° C the degree of dissociation and hydrogen ion concentration of (a) a 0.05 molar solution of benzoic acid with (b) a 0.05 molar solution of benzoic acid containing 0.1 molar sodium benzoate.

12. Calculate the sulfate ion concentration in (a) 0.04 molar H_2SO_4, (b) 0.10 molar $NaHSO_4$, (c) a solution 0.02 molar in H_2SO_4 and 0.02 molar in $NaHSO_4$.
Ans. (a) 0.0011; (b) 0.010; (c) 0.0021 moles/liter

13. Calculate the H^+ and OH^- concentrations at 25° C in each of the following solutions: (a) 0.001 molar H_2SO_4, (b) 0.001 molar $NaHSO_4$, (c) 0.01 molar NH_4OH.

14. Give the formulas of the acids conjugate to (a) methyl alcohol, (b) aniline, (c) dimethyl ether, and of the bases conjugate to (a) methyl alcohol, and (b) phenol.

15. Calculate the hydrolytic constants for each of the following salts at 25° C: Urea hydrochloride, ammonium carbonate, disodium acid phosphate, and sodium bicarbonate.

16. Calculate the degree of hydrolysis and the OH^- ion concentration in (a) a 0.01 molar Na_2CO_3 solution, and (b) a 0.5 molar KCN solution.

17. To what extent will a 0.1 molar solution of aniline acetate ($C_6H_5NH_3C_2H_3O_2$) hydrolyze at 25° C? What will be the pH? *Ans.* 52.8%; pH = 4.71

18. Using the Debye-Hückel limiting law calculate the hydrogen ion activity and pH of a 0.5 molar CH_3COOH solution. How do these values compare with those which would have been obtained on the assumption that ionic concentration is equal to ionic activity?

19. Calculate the pH of the following solutions at 25° C before and after the addition of 1 cc of 0.1 N HCl: (a) 100 cc pure H_2O, (b) 100 cc of 0.1 molar Na_2HPO_4.

20. If the final volume is to be 1 liter, how many moles of HCl will have to be added to 500 cc of 0.1 molar Na_2CO_3 in order to adjust the pH to 7.0? *Ans.* 0.050

21. Assuming no volume change during the titration, calculate the pH after the removal of the first hydrogen when 0.01 molar orthophosphoric acid (H_3PO_4) is titrated with a strong base. What is the pH after the removal of the second hydrogen?

22. Calculate the amounts of 0.1 normal NaOH and pure H_2O which would have to be used in order to make 1-liter portions of buffers having pH's: 6.0, 6.3, and 6.6, starting with 500 cc of 0.1 N NH_4Cl in each case.

23. What pH will result when the following solutions are mixed and made up to a total volume of 1 liter: (a) 500 cc of 0.1 molar KH_2PO_4 + 100 cc of 0.1 N NaOH, (b) 500 cc of 0.1 molar K_2HPO_4 + 50 cc of 0.1 N HCl, (c) 500 cc of 0.01 N benzoic acid + 100 cc of 0.02 N sodium benzoate?

24. From the value of K_{In} for bromcresol green, calculate the pH range over which this indicator is sensitive, and compare your answer with the observed range.

25. The dissociation constant of the complex ion $Ag(CN)_2^-$ is 4×10^{-19}. Calculate the silver ion concentration in a solution which was originally 0.1 molar in KCN and 0.03 molar in $AgNO_3$. *Ans.* 7.5×10^{-18} moles/liter

26. Formulate equilibrium constants for each of the following equilibria:

(a) $AgBr(s) + 2 NH_4OH = Ag(NH_3)_2^+ + Br^- + 2 H_2O$

(b) $IO_3^- + I^- + 6 H^+ = 3 H_2O(l) + I_2(s)$

(c) $H^+ + HCO_3^- = H_2O(l) + CO_2(g)$

27. Calculate the solubility of PbI_2 in (a) pure H_2O, (b) 0.04 molar KI, and (c) 0.04 molar $Pb(NO_3)_2$ at 25° C.

28. The solubility of CaF_2 in water at 18° C is 2.04×10^{-4} moles/liter. Calculate (a) the solubility product, and (b) the solubility in 0.01 molar NaF solution.

    Ans. (a) 3.4×10^{-11}; (b) 3.4×10^{-7} moles/liter

29. At 25° C the solubility product of $AgBrO_3$ is 5.77×10^{-5}. Using the Debye-Hückel limiting law, calculate its solubility in (a) pure H_2O, and (b) 0.01 molar $KBrO_3$. *Ans.* (a) 0.0084, (b) 0.0051 mole/liter

30. At 25° C the solubility product of FeC_2O_4 is 2.1×10^{-7}. Using the Debye-Hückel limiting law, calculate the solubility of FeC_2O_4 in (a) a solution 0.002 molar in $MgSO_4$ and 0.005 molar in KNO_3, (b) pure H_2O.

31. The solubility product of CaF_2 at 18° C is 3.4×10^{-11} while that of $CaCO_3$ is 9.5×10^{-9}. What will be the nature of the first precipitate when a solution of $CaCl_2$ is added to a solution which is 0.05 molar in NaF and 0.02 molar in Na_2CO_3? In a 0.02 molar solution of Na_2CO_3 what is the minimum concentration of NaF at which CaF_2 and $CaCO_3$ will precipitate simultaneously?

32. Using data given in problem 25, predict whether or not AgCl would be precipitated from a solution which is 0.02 molar in NaCl and 0.05 molar in $KAg(CN)_2$.

33. The solubility product of PbI_2 is 7.47×10^{-9} at 15° C and 1.39×10^{-8} at 25° C. Calculate (a) the molar heat of solution of PbI_2, (b) the solubility in moles per liter at 75° C, and (c) compare the latter result with the solubility at 15° C.

    Ans. (a) 1060 cal/mole; (b) 0.00357; (c) 0.00123 mole/liter

The Electromotive Force of Cells

Introduction. Electrochemical cells may be used to perform two functions, namely, (1) *to convert chemical energy into electrical,* and (2) *to convert electrical energy into chemical.* In the common dry cell and the lead storage battery we have converters of chemical into electrical energy, while in the charging of the storage battery and in the electrolytic purification of copper electrical energy is used to bring about chemical action. We shall divide our discussion of cells along these functional lines, and consider in this chapter cells as sources of electrical energy, while in the following chapter certain phenomena associated with the passage of electricity through cells will be elaborated on.

Before proceeding, a distinction should be made between the terms *cell* and *battery,* since they are not synonymous. A cell is a single arrangement of two electrodes and an electrolyte capable of yielding electricity due to chemical action within the cell, or of producing chemical action due to passage of electricity through the cell. A battery, on the other hand, is a combination of two or more cells arranged in series or parallel for the performance of either one or the other of these tasks. Thus, the ordinary 6-volt lead storage battery is a combination of three 2-volt cells connected in series.

Reversible and Irreversible Cells. In dealing with the energy relations of cells thermodynamic principles find very extensive application. However, the use of these principles is subject to one very important restriction, namely, that the processes to which the principles are applied be *reversible.* It will be recalled that the conditions for the thermodynamic reversibility of processes are, (1) that the driving and opposing forces be only infinitesimally different from each other, and (2) that it should be possible to reverse any change taking place by applying a force infinitesimally greater than the one acting. When these requirements are satisfied by a cell, the cell is *reversible,* and its potential difference measured under appropriate conditions may be substituted into the relevant thermodynamic relations. When these conditions are not satisfied, the cell is said to be *irreversible,* and the thermodynamic equations do not apply.

The difference between reversible and irreversible cells may be illustrated with the following two examples. Consider first a cell composed of a zinc and a silver-silver chloride electrode, both dipping into a solution of zinc chloride. When the two electrodes are connected externally through a conductor, electrons flow through the outer circuit from the zinc to the silver-silver chloride. During this passage of current zinc dissolves at one electrode to form zinc ions, while at the other electrode the reaction

$$AgCl(s) = Ag(s) + Cl^- \tag{1}$$
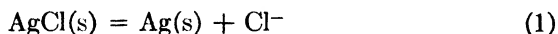
takes place. The net reaction for the cell is, therefore,

$$\frac{1}{2} Zn(s) + AgCl(s) = Ag(s) + \frac{1}{2} Zn^{++} + Cl^- \tag{2}$$
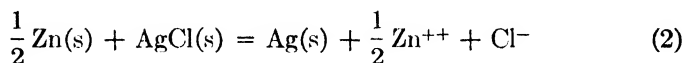
and this process continues as long as the external opposing potential is infinitesimally smaller than that of the cell. However, as soon as the opposing potential becomes ever so slightly larger than that of the cell, the direction of current flow is reversed, and so is the cell reaction. In other words, zinc ions go to form zinc at one electrode, silver chloride is formed from silver and chloride ions at the other, and the overall cell reaction becomes

$$Ag(s) + \frac{1}{2} Zn^{++} + Cl^- = \frac{1}{2} Zn(s) + AgCl(s) \tag{3}$$
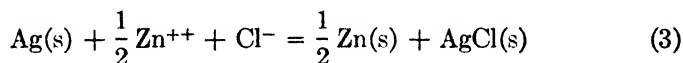
From this description it is obvious that the cell in question meets the second condition of reversibility. Again, the first condition can be satisfied by drawing from or passing through the cell a *very minute* current. Hence, this cell is reversible, and may be treated by thermodynamic methods.

Consider now instead a cell composed of zinc and silver electrodes immersed in a solution of sulfuric acid. When the two electrodes are short circuited, zinc dissolves with evolution of hydrogen to form zinc sulfate according to the scheme

$$Zn(s) + H_2SO_4 = ZnSO_4 + H_2(g) \tag{4}$$
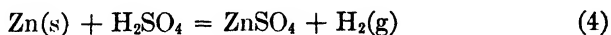
However, when the cell is connected with an external source of potential slightly greater than its own, silver dissolves at one electrode, hydrogen is evolved at the other, and the cell reaction becomes

$$2 Ag(s) + H_2SO_4 = Ag_2SO_4 + H_2(g) \tag{5}$$
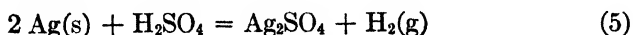
From equations (4) and (5) it is readily evident that, even though this cell may be made to satisfy the first condition of reversibility, the second does not hold, and consequently the cell cannot be reversible. To such a cell thermodynamic principles will not apply with rigidity.

There are a number of other types of irreversibility to which reference will be made later. For the present, suffice it to point out that in theoretical study of the potentials of cells it is the reversible ones which are of importance, and it is to these that we shall confine our prime attention.

Electromotive Force and Its Measurement. If a cell is connected in series with a galvanometer and the circuit is closed, the galvanometer is deflected, indicating that a current is flowing through the circuit. The fact that a current passes from one electrode to the other is evidence for the existence of a potential difference between the electrodes, for without the presence of a potential difference no electricity can flow from one point to another. *This difference of potential which causes a current to flow from the electrode of higher potential to the one of lower is called the electromotive force,* abbreviated emf, *of the cell,* and is expressed in volts.

Although comparatively little is known about the nature of the electromotive force, it can be measured and handled in equations. The most common method of determining the potential difference between any two points in an electric circuit is to connect a voltmeter *across* the two points. The potential difference or voltage is read then directly from the instrument. A serious objection to the use of the voltmeter for accurate measurement of cell potentials is that it draws some current from the cell, causing thereby a change in the emf of the cell due to formation of reaction products at the electrodes and changes in the concentration of the electrolyte around the electrodes. Again, with an appreciable current flowing, part of the emf will have to be utilized to overcome the internal resistance of the cell, with the result that the potential measured on the voltmeter will constitute only part of the total emf of the cell. For these reasons precise emf's of cells are never determined with voltmeters. Instead *potentiometers* are used which require extremely small currents at balance.

Potentiometers for emf measurements operate on the Poggendorff compensation principle. In this method the unknown emf is opposed by another known emf until the two are equal as shown by no deflection
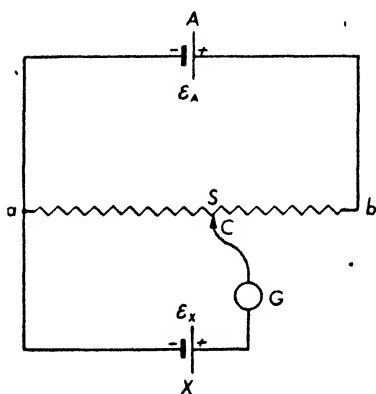
Fig. 1. Poggendorff Compensation Method for Measuring Emf

on a galvanometer present in the circuit. The setup and the conditions at balance may be understood from the diagram shown in Fig. 1. In this diagram A is a cell of known emf $\mathcal{E}_A$, whose potential is impressed

across a uniform resistance ab. Connected with A in such a way that the two emf's oppose each other is the source X of unknown potential $\mathcal{E}_X$. To find $\mathcal{E}_X$ the sliding contact C is moved along ab until a position S is found at which the galvanometer G gives no deflection. From $\mathcal{E}_A$, and the distances ab and aS, the unknown emf $\mathcal{E}_X$ is found as follows. Since $\mathcal{E}_A$ is impressed across the full length ab, $\mathcal{E}_A$ for any given current passing through the resistance must be proportional to ab. Again, as $\mathcal{E}_X$ is impressed only across the distance aS, it must be proportional to this length. Consequently, on dividing $\mathcal{E}_X$ by $\mathcal{E}_A$ we obtain

$$\frac{\mathcal{E}_X}{\mathcal{E}_A} = \frac{aS}{ab}$$

and therefore,

$$\mathcal{E}_X = \left(\frac{aS}{ab}\right)\mathcal{E}_A \tag{6}$$

The only requirements that need be met here are that $\mathcal{E}_A$ be larger than $\mathcal{E}_X$, that the wire ab be uniform, and that the galvanometer be sufficiently sensitive to allow a balance of the potentials without appreciable current flow. Once these requirements are satisfied, the unknown emf can be obtained through equation (6) under conditions where the cell suffers no disturbance due to passage of current, and hence under conditions approximating very closely true reversibility.

To permit *direct reading* of voltage and to conserve the standard source of potential A, the setup shown in Fig. 1 is modified in practice to that given in Fig. 2. In this diagram W is the *working cell* whose emf can be impressed across ab, which is calibrated in volts, through the variable resistance R. Against this cell may be applied either the unknown emf X or the standard cell S.C. through the double-pole double-throw switch D. Before the emf of X can be measured, the slide wire must be standardized to read emf direct. This is done as follows. First pointer C is set at a point S' along ab corresponding to the value of the emf of the standard cell, say 1.0183 volts. Next the switch D is thrown to the S.C. side, the tapping key is closed gently, and F is moved along the variable resistance R until the galvanometer G shows no deflection. When this point of balance is established, the current flowing through ab is of such a magnitude as to make the potential drop between a and S' exactly 1.0183 volts, and hence the current is just right to make the voltage drop anywhere along ab correspond to the voltage markings on the slide wire. In other words, the potentiometer has been standardized to read voltage directly.

Once this operation has been completed, R is left undisturbed, the switch D is shifted to the unknown emf side, the key is closed, and this time C is moved along ab until a point S is found at which the galvanom-

eter G again shows no deflection. The reading of the slide wire at S gives then the voltage of X directly.

In laboratory potentiometers the slide wire ab is usually not a single unit, but consists first of a series of coils of nominal 0.1 volt values, and
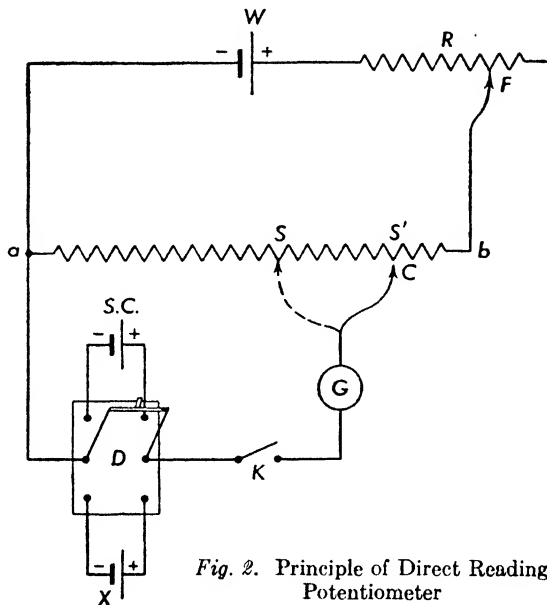
Fig. 2. Principle of Direct Reading
Potentiometer

an extended slide wire covering in toto a 0.1 volt range. By this means the Leeds and Northrup Student Type Potentiometer can be used to measure potentials up to 1.6 volts to 1×10^{-4} volt. In more precise potentiometers, such as the Leeds and Northrup Type K–2, the precision can be extended to 1×10^{-5} volt, and for small values of emf down to 1×10^{-7} volt. With the latter instrument galvanometers of very high sensitivity are required.

Standard Cells. The accuracy with which emf measurements can be made by the potentiometric method depends to a large degree on the accuracy with which the potentials of the cells used for reference are known. It is essential, therefore, to have available for comparison cells whose potentials are reproducible, constant with time, and well known. Also, such cells should be reversible, should not be subject to permanent damage due to passage of current through them, and should preferably have low temperature coefficients of emf. The cells that most closely approximate these properties are either the *Weston unsaturated standard cell* or the *Weston saturated standard cell*. At one time the Clark standard cell was also used, but because of its high temperature coefficient it has been almost completely superseded by the Weston cells.

The usual form of the Weston saturated standard cell is illustrated in Fig. 3. The cell consists of an H-shaped glass vessel containing in each arm one of the electrodes and filled throughout with the electrolyte. Contact with the active material is made through short platinum wires sealed into the bottoms of the arms. The positive electrode consists of mercury covered with a paste of mercurous sulfate and mercury. The negative electrode, in turn, consists of a cadmium amalgam containing 12 to 14 per cent cadmium by weight. Over both electrodes are sprinkled some crystals of solid $CdSO_4 \cdot 8/3\ H_2O$, the entire cell is filled with a saturated solution of cadmium sulfate, and the cell is closed with

Fig. 3. Saturated Weston Standard Cell

either corks and sealing wax, or the arms are drawn off. The purpose of the solid crystals of $CdSO_4 \cdot 8/3\ H_2O$ is to keep the electrolyte saturated with this phase at all temperatures.

The operation of the Weston saturated standard cell depends on the reversible reaction

$$Cd(s) + Hg_2SO_4(s) \rightleftharpoons CdSO_4(s) + 2\ Hg(l) \qquad (7)$$

When acting as a source of current, cadmium dissolves to form cadmium sulfate at the negative electrode, while mercury is deposited from the mercurous sulfate at the positive electrode. Since the solution is saturated, an amount of cadmium equivalent to that dissolved will combine with SO_4^{--} ions to precipitate cadmium sulfate as $CdSO_4 \cdot 8/3\ H_2O$. On the other hand, by depositing mercury the solution becomes unsaturated with respect to mercurous sulfate in the immediate vicinity of the electrode, and hence some mercurous sulfate dissolves until the solution is again saturated with the latter. However, when the cell is subjected to a potential higher than its own, mercury dissolves at one electrode, and cadmium is deposited at the other. At the same time some mercurous sulfate precipitates from solution, while some $CdSO_4 \cdot 8/3\ H_2O$ dissolves until the solution again becomes saturated with respect to the two phases. In this manner the solution stays saturated at all times, and preserves thus a fixed concentration of mercurous and cadmium ions in contact with the electrodes.

In construction the *unsaturated* Weston cell is similar to the saturated, except that the crystals of solid $CdSO_4 \cdot 8/3\ H_2O$ are omitted, and the electrolyte is a solution of cadmium sulfate *saturated only at 4° C.* At

all temperatures above 4° C, the electrolyte in this cell is, therefore, unsaturated. Otherwise the reaction responsible for the emf is the same as in the saturated cell, and so is the mechanism of cell operation.

The Weston standard cells when carefully prepared from specially purified materials and not subjected to abuse will preserve their potentials excellently for many years. The Weston *saturated cell*, or as it is frequently called the Weston *normal cell*, is the working standard of emf adopted by the United States Bureau of Standards. Its potential at any temperature $t°$ C is given by the equation

$$\mathcal{E}_t = 1.01830 - 4.06 \times 10^{-5}(t - 20) - 9.5 \times 10^{-7}(t - 20)^2 \quad (8)$$

From equation (8) the potential of the cell at 20° C follows as 1.01830 international volts, while at 25° C it is 1.01807 volts. The unsaturated cell is employed only as a secondary standard, possesses a lower temperature coefficient than the saturated, and has a potential close to 1.0186 volts at 20° C which may be assumed to be constant for any measurements around room temperature.

Cell Reaction and Emf. In studying cells it is necessary to determine not only the emf of the cell but also the reaction responsible for producing the emf. The nature of the reaction proceeding in a cell cannot be deduced from the emf alone. However, it can be ascertained from the *direction of the current flow*, i.e., electron flow, during the measurement of the emf. Consequently, in measuring the emf of a cell by the potentiometric method it is essential to record not only the emf of the cell, but also *the polarity of the electrodes.* The latter are determined by observing the manner in which the electrodes must be connected to the standard cell in order to obtain a balance of the potentiometer. Since such a balance is possible only when the cells are connected so as to oppose each other, it must follow that the *electrode connected to the negative side of the standard cell is the negative electrode*, while the one connected to the positive side is the positive electrode. And, since a negative electrode contains electrons in excess of the number present on the positive electrode, *electrons will have to flow from the negative to the positive electrode in the external circuit.*

Once the direction of current flow is known, the processes responsible for emission of electrons at the negative electrode and the taking on of the electrons at the positive electrode can be ascertained from the nature of the electrode materials on the basis of our ordinary concepts of oxidation and reduction. The student is aware that any reaction that gives off electrons must be an oxidation, while any reaction that involves a taking on of electrons with consequent decrease in positive valence must be a reduction. Hence, *an oxidation must occur at the negative electrode* where electrons are given off, and *a reduction must take place at the*

positive electrode where electrons enter. By *adding* the oxidation reaction at the negative electrode to the reduction at the positive electrode we obtain the cell reaction.

This manner of arriving at the cell reaction from the direction of electron flow may be illustrated with an example. Consider a cell composed of a zinc electrode dipping in a solution of zinc ions at unit activity, and a cadmium electrode dipping into a solution of cadmium ions at the same activity; in other words, the cell

$$\text{Zn} \mid \text{Zn}^{++}(a = 1) \parallel \text{Cd}^{++}(a = 1) \mid \text{Cd} \tag{9}$$
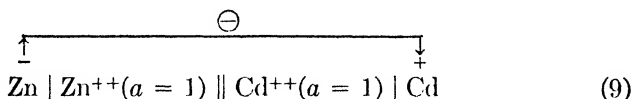
At 25° C it is found with the aid of a potentiometer that this cell has an emf of $\mathcal{E}_{25°} = 0.3590$ volt, and that the zinc electrode is negative. Hence the electron flow through the external circuit must be from the zinc to the cadmium electrodes, and consequently an oxidation occurs at the zinc electrode, a reduction at the cadmium electrode. The only oxidation process possible at the zinc electrode is, obviously,

$$\text{Zn(s)} = \text{Zn}^{++}(a = 1) + 2 \ominus$$
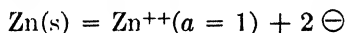
Again, the reduction at the cadmium electrode is

$$\text{Cd}^{++}(a = 1) + 2 \ominus = \text{Cd(s)}$$
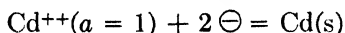
The overall cell reaction must be, then,

$$\text{Zn(s)} + \text{Cd}^{++}(a = 1) = \text{Zn}^{++}(a = 1) + \text{Cd(s)} \tag{10}$$
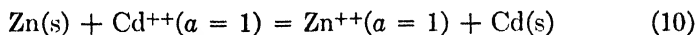
and this reaction yields an emf of $\mathcal{E} = 0.3590$ volt at 25° C.

It is very important to realize that an emf without the reaction responsible for it is as meaningless as the age of an individual without specifying which individual is being referred to. For absolute clarity each emf must be accompanied by the reaction to which it refers, as well as by a complete statement of the nature of the phases, their concentrations, and the temperature.

Convention Regarding Sign of Emf. The net electrical work performed by a reaction yielding an emf $\mathcal{E}$ and supplying a quantity of electricity Q is $Q\mathcal{E}$. For each equivalent reacting Q is equal to the value of the faraday $\mathcal{F}$, and hence for n equivalents reacting $Q = n\mathcal{F}$. Therefore, the electrical work obtained from any reaction supplying $n\mathcal{F}$ coulombs of electricity at a potential $\mathcal{E}$ is

$$\text{Net electrical work} = n\mathcal{F}\mathcal{E}$$

But any work performed by a cell can be accomplished only at the expense of a *decrease in free energy* occurring within the cell due to

chemical reaction. Further, when the electrical work is a *maximum*, as when the cell operates *reversibly*, the decrease in free energy $-\Delta F$ must equal the electrical work done. We obtain, therefore, the equality

$$\Delta F = - n\mathcal{F}\mathcal{E} \tag{11}$$

from which we see that the reversible emf of any cell is determined by the free energy change of the reaction going on in the cell. Equation (11) is the "bridge" between thermodynamics and electrochemistry. Through it the evaluation of various thermodynamic properties of reactions becomes possible from emf measurements.

It was shown in Chapter XI that for any spontaneous reaction ΔF is negative, for any nonspontaneous reaction ΔF is positive, while for any reaction in equilibrium $\Delta F = 0$. In view of this, it may be deduced from equation (11) that for any spontaneous reaction $\mathcal{E}$ will have to be positive, for any nonspontaneous reaction $\mathcal{E}$ will be negative, while for any reaction in equilibrium $\mathcal{E}$ will have to be zero. These relations between the spontaneity of a reaction and the signs of ΔF and $\mathcal{E}$ are summarized in Table 1. They indicate that in order to have conformity between emf and ΔF it is necessary to prefix each and every emf with the appropriate sign depending on whether the reaction as written is the spontaneous one.

TABLE 1

RELATION BETWEEN SIGNS OF ΔF AND $\mathcal{E}$

Reaction	ΔF	$\mathcal{E}$
Spontaneous	−	+
Nonspontaneous	+	−
Equilibrium	0	0

The emf's of all cells under discussion here result from the spontaneous reactions occurring within the cells, and hence for all of them $\mathcal{E}$ is positive. However, it is not always possible to tell *a priori* which of two processes possible, namely, the forward and reverse reactions, is the spontaneous one. Consequently it is necessary to have some means of ascertaining which reaction results in the measured positive emf. For this purpose we lay down the following rule: *If a cell is written on paper such that the negative electrode is on the left and the positive electrode on the right,* in other words, *so that electrons flow from left to right through the external circuit, the reaction deduced on the basis of oxidation-reduction described above is the spontaneous one, and the emf of the cell is positive.* By observing this rule the spontaneous reaction in any cell may be arrived at, and the correct sign ascribed to the emf corresponding to it.

We may illustrate the application of this rule with the cell given in equation (9). From the equation we see that the cell is written in

accordance with the rule, i.e., the negative electrode on the left, and hence the reaction arrived at above, equation (10), is the spontaneous one. For this reaction, then, $\mathcal{E}_{25°} = + 0.3590$ volt. It follows, therefore, that the reverse of (10), namely,

$$Cd(s) + Zn^{++}(a = 1) = Zn(s) + Cd^{++}(a = 1) \qquad (12)$$
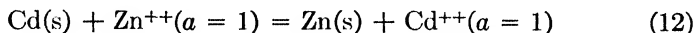
is not spontaneous, and for this reaction the emf at 25° C will be $\mathcal{E}_{25}° = - 0.3590$. This means that in order to carry out the reaction in equation (12) at least 0.3590 volt will have to be applied to the cell; and, instead of obtaining electrical work from the cell, work will have to be done to get the reaction to go.

Single Electrode Potentials. Any electrochemical reaction is the sum of the two electrode reactions, of which one involves a loss of electrons and hence an oxidation, the other a removal of electrons from the electrode or a reduction. Similarly every cell emf may be thought of as being composed of two individual electrode potentials, called *single electrode potentials*, such that their *algebraic sum* is equal to the total emf of the cell. With these single electrode potentials the emf of any cell readily follows as the sum of the electrode potentials, just as the overall cell reaction is the sum of the single electrode reactions.

Although the emf of any cell may be determined in absolute units, namely, volts, for single electrode potentials we must employ an arbitrary standard electrode to which all other electrode potentials are referred. Many attempts have been made to prepare a reference electrode possessing absolute zero difference in potential between itself and the solution around it, but so far these attempts have not been very successful and there is no generally accepted electrode of true zero potential. For this reason *we define as the reference of emf the standard hydrogen electrode whose potential at all temperatures is taken as zero.* Suffice it to say now that a standard hydrogen electrode consists of a piece of platinized platinum foil surrounded by hydrogen gas at 1 atmosphere pressure, and immersed in a solution containing hydrogen ions at unit activity. Details on the theory and operation of this electrode will be given later.

Because of the many experimental difficulties involved in the preparation and use of the standard hydrogen electrode, secondary reference electrodes have been compared with the former, and are widely used. Among these are the three *calomel electrodes*, whose emf depends on the spontaneous process

$$Hg_2Cl_2(s) = 2 Hg(l) + 2 Cl^-(C = x) \qquad (13)$$
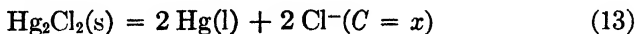
where x is 0.1 N and 1 N potassium chloride for the 0.1 N and 1 N calomel electrodes, and a saturated potassium chloride solution for the

saturated calomel electrode. A common form in which these electrodes are encountered in the laboratory is shown in Fig. 4. The electrode consists of a glass vessel A to which are attached the sidearm B for making electrical contact with the electrode, and the arm C for inser-

Fig. 4. Calomel Electrode

tion in any desired solution. Into the bottom of A is sealed a platinum wire over which are placed in turn a layer of specially purified mercury, a paste of mercury and calomel, and then the appropriate potassium chloride solution saturated with calomel so as to fill the cell and the arm C. In the 0.1 N calomel this solution is 0.1 N potassium chloride, in the 1 N calomel it is 1 N potassium chloride, and in the saturated calomel it is an aqueous solution saturated with both potassium chloride and mercurous chloride. In the latter electrode some crystals of potassium chloride are also placed over the mercury-mercurous chloride paste to keep the solution saturated at all times.
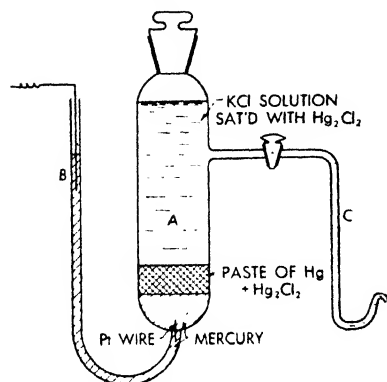
The potentials of the various calomel electrodes depend at each temperature on the concentration of potassium chloride with which the electrode vessels are filled. In Table 2 are given the potentials of these electrodes as a function of temperature, the emf at 25° C, and the corresponding cell reactions. These emf's are highest for the 0.1 N calomel and decrease as the concentration is increased. Further, the 0.1 N calomel has the smallest temperature coefficient of emf.

TABLE 2

POTENTIALS OF CALOMEL ELECTRODES [1]

Electrode	Symbol	Emf	$\mathcal{E}_{25°C}$	Reaction
0.1 N calomel	$Hg \mid Hg_2Cl_2(s)$, KCl(0.1 N)	$\mathcal{E} = 0.3338 - 7 \times 10^{-5}(t - 25)$	0.3338	$Hg_2Cl_2(s) = 2\,Hg\,(l) + 2\,Cl^-(0.1\,N)$
1 N calomel	$Hg \mid Hg_2Cl_2(s)$, KCl(1 N)	$\mathcal{E} = 0.2800 - 2.4 \times 10^{-4}(t - 25)$	0.2800	$Hg_2Cl_2(s) = 2\,Hg\,(l) + 2\,Cl^-(1\,N)$
Saturated calomel	$Hg \mid Hg_2Cl_2(s)$, KCl (sat'd)	$\mathcal{E} = 0.2415 - 7.6 \times 10^{-4}(t - 25)$	0.2415	$Hg_2Cl_2(s) = 2\,Hg\,(l) + 2\,Cl^-$ (sat'd)

[1] Hamer, Trans. Electrochem. Soc., **72**, 45 (1937).

The manner in which these auxiliary reference electrodes are combined with other electrodes to form cells is shown in Fig. 5. Here A is the reference calomel electrode, B the other single electrode whose potential is to be determined, and C a *salt bridge* to permit electrical contact
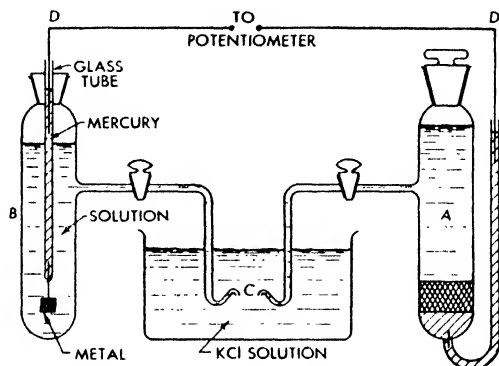
Fig. 5. Determination of Single Electrode Potentials

between A and B. The salt bridge consists usually of a beaker filled with 1 N or saturated potassium chloride solution into which the side-arms of the electrodes are immersed. Electrical contact with the poten-tiometer is made by means of the wires D and D'.

Calculation of Single Electrode Potentials. From emf's of cells involving various electrodes in combination with reference electrodes, single electrode potentials are readily calculated. To illustrate the pro-cedure consider first the cell

$$\overline{Cd} \mid Cd^{++}(a = 1) \parallel KCl(1\ N),\ Hg_2Cl_2(s) \mid Hg \qquad (14)$$
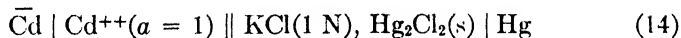
consisting of a cadmium-cadmium ion electrode and a 1 N calomel. For this cell it is found that at 25° C $\mathcal{E} = 0.6821$ volt, and that the cadmium electrode is negative. Consequently the cadmium electrode undergoes the oxidation reaction

$$Cd(s) = Cd^{++}(a = 1) + 2\ominus$$
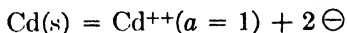
while the calomel electrode undergoes the reduction,

$$Hg_2Cl_2(s) + 2\ominus = 2\ Hg(l) + 2\ Cl^-(1\ N)$$
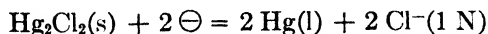
For the latter process Table 2 shows that the emf at 25° C is $\mathcal{E}_c = +0.2800$ volt. Consequently, if we let $\mathcal{E}_{Cd}$ be the single electrode potential of the cadmium electrode,

$$\mathcal{E} = \mathcal{E}_{Cd} + \mathcal{E}_c$$
$$0.6821 = \mathcal{E}_{Cd} + 0.2800$$
and $$\mathcal{E}_{Cd} = 0.4021\ \text{volt}$$

Hence the spontaneous reaction and the corresponding emf at 25° C for the $Cd \mid Cd^{++}(a = 1)$ electrode are

$$Cd \mid Cd^{++}(a = 1) : \quad Cd(s) = Cd^{++}(a = 1) + 2 \ominus \quad \mathcal{E}_{25°} = 0.4021 \text{ volt}$$
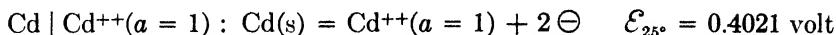
As a second example let us take the cell

$$Hg \mid Hg_2Cl_2(s), KCl(1 N) \parallel Cu^{++}(a = 1) \mid Cu \qquad (15)$$
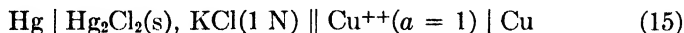
consisting of a 1 N calomel and a copper electrode dipping in a solution of cupric ions at unit activity. At 25° C the emf of this cell is $\mathcal{E} = 0.0648$ volt, and *the calomel electrode is negative.* Since the calomel electrode is negative, the reaction occurring in it must be an oxidation, or

$$2 Hg(l) + 2 Cl^-(1 N) = Hg_2Cl_2(s) + 2 \ominus$$
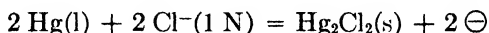
Further, as the reduction reaction yields an emf of $+ 0.2800$ volt, on oxidation the emf of the calomel in this cell must be $\mathcal{E}_C = - 0.2800$ volt. The corresponding reduction at the copper electrode must involve the reaction

$$Cu^{++}(a = 1) + 2 \ominus = Cu(s)$$
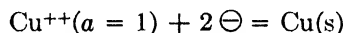
with a single electrode potential $\mathcal{E}_{Cu}$. We obtain, then,

$$\mathcal{E} = \mathcal{E}_C + \mathcal{E}_{Cu}$$
$$0.0648 = - 0.2800 + \mathcal{E}_{Cu}$$
$$\mathcal{E}_{Cu} = 0.3448 \text{ volt}$$

and therefore,

$$Cu \mid Cu^{++}(a = 1) : \quad Cu^{++}(a = 1) + 2 \ominus = Cu(s) \quad \mathcal{E}_{25°} = 0.3448 \text{ volt}$$
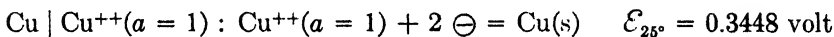
It will be observed that whereas for the cadmium electrode the oxidation is the spontaneous process, for the copper electrode the reverse is true, namely, the reduction of copper ions to copper proceeds spontaneously.

Calculation of Cell Emf's from Single Electrode Potentials. When the single electrode potentials and reactions are available, the calculations may be reversed to predict cell emf's and cell reactions. Suppose it is desired to know the reaction and the emf at 25° C for the cell

$$Cd \mid Cd^{++}(a = 1) \parallel Cu^{++}(a = 1) \mid Cu \qquad (16)$$
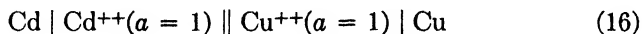
Let us assume that the cell as written is correct, i.e., that the cadmium electrode is negative. Then, for the oxidation at the cadmium electrode, we have

$$Cd(s) = Cd^{++}(a = 1) + 2 \ominus \qquad \mathcal{E}_{25°} = 0.4021 \text{ volt}$$
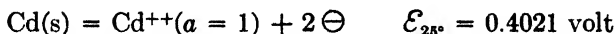
Again, for the reduction at the copper electrode, we have

$$Cu^{++}(a = 1) + 2 \ominus = Cu(s) \qquad \mathcal{E}_{25°} = 0.3448 \text{ volt}$$
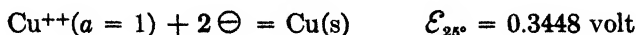
Adding now the single electrode reactions and the single electrode potentials we find, for the cell reaction,

$$Cd(s) + Cu^{++}(a = 1) = Cd^{++}(a = 1) + Cu(s) \qquad (17a)$$
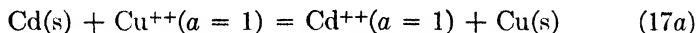
and for the emf,

$$\mathcal{E}_{cell} = \mathcal{E}_{Cd} + \mathcal{E}_{Cu}$$
$$= 0.4021 + 0.3448$$
$$= + 0.7469 \text{ volt at } 25° \text{ C} \qquad (17b)$$

Since the calculated emf is positive, the assumption made with respect to the polarity of the electrodes is correct, and the cell reaction as given is the spontaneous one.

On the other hand, had we assumed the copper electrode in equation (16) to be negative, the latter would have to undergo the oxidation

$$Cu(s) = Cu^{++}(a = 1) + 2 \ominus$$
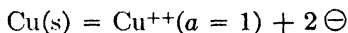
with emf equal to $\mathcal{E}_{Cu} = - 0.3448$ volt; the cadmium electrode would have to suffer the reduction

$$Cd^{++}(a = 1) + 2 \ominus = Cd(s)$$

with emf $\mathcal{E}_{Cd} = - 0.4021$ volt; and we would have found for the cell reaction and emf

$$Cu(s) + Cd^{++}(a = 1) = Cu^{++}(a = 1) + Cd(s)$$
$$\mathcal{E}_{cell} = - 0.7469 \text{ volt at } 25° \text{ C} \qquad (18)$$
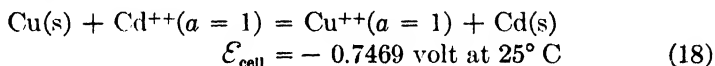
Inspection of equation (18) reveals that since $\mathcal{E}_{cell}$ is negative the reaction as written is not spontaneous, and consequently the wrong assumption was made with respect to the polarity of the electrodes. All that need be done to rectify the situation is *to reverse the reaction* written in equation (18) and *to change the sign of* $\mathcal{E}_{cell}$. The result is then the spontaneous reaction and emf of the cell as given in equations (17a) and (17b).

Summary of Rules. Before proceeding it may not be out of place to summarize here the various principles we have employed above with respect to cell reactions and emf's. These are:

(1) Any cell reaction is the *sum* of the single electrode reactions *as they occur in the cell.*

(a) At the *negative electrode* the reaction is an *oxidation.*

(b) At the *positive electrode* the reaction is a *reduction.*

(2) The total cell emf is the algebraic sum of the single electrode potentials provided each emf is affixed with the sign corresponding to the spontaneity of the reaction taking place at the electrode.

(a) For the *spontaneous* electrode reaction the single electrode potential is *positive* with respect to standard hydrogen electrode.

(b) For the *nonspontaneous* electrode reaction the single electrode potential is *negative* with respect to standard hydrogen electrode.

(c) The same rule with respect to the sign applies also to the cell reaction.

(3) If any cell is written down with the *negative electrode on the left,* so that electrons flow through the external circuit from left to right, the cell reaction deduced by rule (1) will be the *spontaneous* process, and the emf deduced by rule (2) for, the cell will be *positive.*

(4) If the wrong assumption be made with respect to the polarity of the electrodes, rule (3) will yield a *negative* emf corresponding to the *nonspontaneous* reaction. To obtain the spontaneous reaction and its emf, all that need be done is *to reverse the reaction and change the sign of the emf without changing its magnitude.*

By adhering *rigidly* to these rules the student can avoid any confusion in regard to signs or reactions.

Thermodynamics and Emf: ΔH and ΔS from Emf Data. According to equation (12), Chapter XI, the change in free energy ΔF for *any* process is related to the change in heat content, i.e., the heat of reaction at constant pressure, by the Gibbs-Helmholtz equation

$$\Delta F - \Delta H = T\left[\frac{\partial(\Delta F)}{\partial T}\right]_P \qquad (19)$$

Again, the fundamental relation between emf and ΔF is equation (11), namely,

$$\Delta F = -n\mathcal{F}\mathcal{E} \qquad (11)$$

If equation (11) is differentiated with respect to temperature, then,

$$\left[\frac{\partial(\Delta F)}{\partial T}\right]_P = -n\mathcal{F}\left(\frac{\partial\mathcal{E}}{\partial T}\right)_P \qquad (20)$$

and on substitution of equations (20) and (11) into equation (19) we find for ΔH

$$-n\mathcal{F}\mathcal{E} - \Delta H = -n\mathcal{F}T\left(\frac{\partial\mathcal{E}}{\partial T}\right)_P$$

$$\Delta H = -n\mathcal{F}\mathcal{E} + n\mathcal{F}T\left(\frac{\partial\mathcal{E}}{\partial T}\right)_P$$

$$= n\mathcal{F}\left[T\left(\frac{\partial\mathcal{E}}{\partial T}\right)_P - \mathcal{E}\right] \qquad (21)$$

Equation (21) permits the calculation of the heat of a reaction from the measured emf $\mathcal{E}$ and temperature coefficient of emf, $(\partial\mathcal{E}/\partial T)_P$, of the reaction, or $(\partial\mathcal{E}/\partial T)_P$ from ΔH and $\mathcal{E}$. In other words, equation (21) makes it possible to estimate heats of reaction from emf data. In using this equation, ΔH follows in *joules* when $\mathcal{E}$ and $\mathcal{F}$ are in volts and cou-

lombs respectively, while ΔH must be substituted in *joules* to obtain $\mathcal{E}$ in volts, or $(\partial \mathcal{E}/\partial T)_P$ in volts per $°$ K.

To show the use of equation (21) for calculating ΔH, we may take the reaction

$$Zn(s) + 2\,AgCl(s) = ZnCl_2(0.555\ m) + 2\,Ag(s) \tag{22}$$
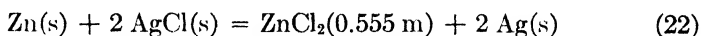
for which $\mathcal{E}_{0° C} = 1.015$ volt while $(\partial \mathcal{E}/\partial T)_P = -4.02 \times 10^{-4}$ volts per degree. Since in equation (22) 2 faradays are necessary to accomplish the change, $n = 2$; and, with $\mathcal{F} = 96{,}500$ and $T = 273.2°$ K, the value of ΔH results as

$$
\begin{aligned}
\Delta H &= n\mathcal{F}\left[T\left(\frac{\partial \mathcal{E}}{\partial T}\right)_P - \mathcal{E} \right] \\
&= 2 \times 95{,}500[273.2(-4.02 \times 10^{-4}) - 1.015] \\
&= -217{,}100 \text{ joules} \\
&= -\frac{217{,}100}{4.185} = -51{,}900 \text{ cal}
\end{aligned}
$$

The heat of reaction found calorimetrically is $\Delta H = -52{,}050$ calories, a value in excellent accord with that obtained from emf data.

It should be noticed that $n = 2$ in equation (22) only for the reaction written. Had the reaction been given as

$$\frac{1}{2}\,Zn(s) + AgCl(s) = \frac{1}{2}\,ZnCl_2(0.555\ m) + Ag(s) \tag{22a}$$
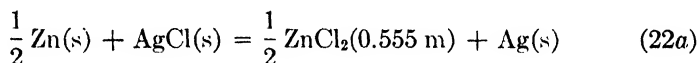
n would be *one*, but $\mathcal{E}$ and $(\partial \mathcal{E}/\partial T)_P$ would remain the same. Under these conditions ΔH would follow as one-half the quantity given above, since only half the amounts of materials would be involved in reaction. For this reason it is essential to state the equation for which ΔH is calculated.

Equation (19) indicates that in general ΔF will differ from ΔH by $T\left[\dfrac{\partial(\Delta F)}{\partial T}\right]_P$, or, in view of equation (20),

$$\Delta F - \Delta H = -n\mathcal{F}T\left(\frac{\partial \mathcal{E}}{\partial T}\right)_P \tag{23}$$

Equation (23) indicates that for any electrochemical reaction ΔF can equal ΔH only when $(\partial \mathcal{E}/\partial T)_P = 0$. However, when the temperature coefficient of a cell is not zero, ΔF will be greater than ΔH for $(\partial \mathcal{E}/\partial T)_P$ negative, and ΔH will be greater than ΔF for $(\partial \mathcal{E}/\partial T)_P$ positive. In fact, it can readily be proved that ΔF and ΔH for any process will be identical only when the entropy change ΔS is zero; otherwise the two will be different. To do this we utilize equation (9), Chapter XI, namely,

$$\Delta F - \Delta H = -T\Delta S \tag{24}$$

On comparison of equations (23) and (24) we find that

$$- T\Delta S = - n\mathcal{J}T \left(\frac{\partial \mathcal{E}}{\partial T}\right)_P$$

and therefore,
$$\Delta S = n\mathcal{J} \left(\frac{\partial \mathcal{E}}{\partial T}\right)_P \tag{25}$$

With the aid of equation (25) the entropy change of a reaction can readily be calculated from the temperature coefficient of emf. For the cell cited in equation (22), $(\partial \mathcal{E}/\partial T)_P = - 4.02 \times 10^{-4}$, $n = 2$. Hence the entropy change of this reaction is

$$\Delta S = 2 \times 96,500(- 4.02 \times 10^{-4})$$
$$= - 77.59 \text{ joules/degree}$$
$$= - 18.55 \text{ cal/degree}$$

Thermodynamics of Electrode Potentials. Single electrode potentials, and through these cell potentials, are determined not only by the nature of the constituents composing the electrodes, but also by the temperature and the activities of the solutions employed. The manner in which the emf's depend on the activities and the temperature is given by the *Nernst equation*, which in turn is deducible from thermodynamics.

For any reaction such as

$$\alpha A + bB \cdots = cC + dD \cdots \tag{26}$$

the change in free energy ΔF, as a function of the *initial* activities of the reactants and the *final* activities of the products, is given by the *reaction isotherm* (Chapter XI, equation (37)) as

$$\Delta F = \Delta F^0 + RT \ln \frac{a_C^c a_D^d \cdots}{a_A^\alpha a_B^b \cdots} \tag{27}$$

Here the a's in the numerator are the activities of the products, those in the denominator the activities of the reactants, and ΔF^0, the standard free energy change, is the free energy change attending the reaction when the activities of products and reactants are all unity. If we substitute into equation (27) $\Delta F = - n\mathcal{J}\mathcal{E}$, and define ΔF^0 by

$$\Delta F^0 = - n\mathcal{J}\mathcal{E}^0 \tag{28}$$

where $\mathcal{E}^0$ is the value of $\mathcal{E}$ corresponding to ΔF^0, equation (27) becomes

$$- n\mathcal{J}\mathcal{E} = - n\mathcal{J}\mathcal{E}^0 + RT \ln \frac{a_C^c a_D^d \cdots}{a_A^\alpha a_B^b \cdots}$$

and hence $\mathcal{E}$ follows as

$$\mathcal{E} = \mathcal{E}^0 - \frac{RT}{n\mathcal{J}} \ln \frac{a_C^c a_D^d \cdots}{a_A^\alpha a_B^b \cdots}$$
$$= \mathcal{E}^0 - \frac{2.3026 \, RT}{n\mathcal{J}} \log_{10} \frac{a_C^c a_D^d}{a_A^\alpha a_B^b} \tag{29}$$

Equation (29) is the *Nernst equation* for the potential of any electrode or cell emf as a function of the initial and final activities. It shows that the emf is determined by the activities of the reacting species, the temperature, and by the quantity $\mathcal{E}^0$. Since ΔF^0 is a constant for any reaction at constant temperature, $\mathcal{E}^0$ *must also be a constant at constant temperature characteristic of the electrode or cell*, and independent of the activities. It is in fact the emf of the electrode or cell when all the activities are unity. $\mathcal{E}^0$ is called either the *standard* or *normal potential of the electrode or cell in question.*

The *standard electrode potentials* at 25° C for a number of electrodes, evaluated by methods to which reference will be made later, are tabulated in Table 3. The corresponding electrode reactions are also given.

TABLE 3

STANDARD ELECTRODE POTENTIALS AT 25° C FOR OXIDATION REACTIONS

Electrode	Electrode Reaction	$\mathcal{E}^0$ (volts)
Li | Li$^+$	Li(s) = Li$^+$ + $\ominus$	+ 3.024
K | K$^+$	K(s) = K$^+$ + $\ominus$	+ 2.9241
Ca | Ca^{++}	Ca(s) = Ca^{++} + 2 $\ominus$	+ 2.763
Na | Na$^+$	Na(s) = Na$^+$ + $\ominus$	+ 2.7146
Zn | Zn^{++}	Zn(s) = Zn^{++} + 2 $\ominus$	+ 0.7611
Fe | Fe^{++}	Fe(s) = Fe^{++} + 2 $\ominus$	+ 0.441
Cd | Cd^{++}	Cd(s) = Cd^{++} + 2 $\ominus$	+ 0.4021
Tl | Tl$^+$	Tl(s) = Tl$^+$ + $\ominus$	+ 0.3385
Ni | Ni^{++}	Ni(s) = Ni^{++} + 2 $\ominus$	+ 0.236
Ag | AgI(s), I$^-$	Ag(s) + I$^-$ = AgI(s) + $\ominus$	+ 0.1522
Sn | Sn^{++}	Sn(s) = Sn^{++} + 2 $\ominus$	+ 0.1406
Pb | Pb^{++}	Pb(s) = Pb^{++} + 2 $\ominus$	+ 0.1264
H$_2$ | H$^+$	H$_2$(g, 1 atm) = 2 H$^+$ + 2 $\ominus$	± 0.0000
Ag | AgBr(s), Br$^-$	Ag(s) + Br$^-$ = AgBr(s) + $\ominus$	− 0.0711
Hg | Hg$_2$Br$_2$(s), Br$^-$	2 Hg(l) + 2 Br$^-$ = Hg$_2$Br$_2$(s) + 2 $\ominus$	− 0.1385
Pt | Sn^{++}, Sn^{++++}	Sn^{++} = Sn^{++++} + 2 $\ominus$	− 0.14
Ag | AgCl(s), Cl$^-$	Ag(s) + Cl$^-$ = AgCl(s) + $\ominus$	− 0.2225
Hg | Hg$_2$Cl$_2$(s), Cl$^-$	2 Hg(l) + 2 Cl$^-$ = Hg$_2$Cl$_2$ + 2 $\ominus$	− 0.2680
Cu | Cu^{++}	Cu(s) = Cu^{++} + 2 $\ominus$	− 0.3441
I$_2$ | I$^-$	2 I$^-$ = I$_2$(s) + 2 $\ominus$	− 0.5362
Hg | Hg$_2$SO$_4$(s), SO$_4^{--}$	2 Hg(l) + SO$_4^{--}$ = Hg$_2$SO$_4$(s) + 2 $\ominus$	− 0.6141
Pt | Fe^{++}, Fe^{+++}	Fe^{++} = Fe^{+++} + $\ominus$	− 0.7477
Ag | Ag$^+$	Ag(s) = Ag$^+$ + $\ominus$	− 0.7990
Br$_2$ | Br$^-$	2 Br$^-$ = Br$_2$(l) + 2 $\ominus$	− 1.0659
Pt | Tl$^+$, Tl^{+++}	Tl$^+$ = Tl^{+++} + 2 $\ominus$	− 1.211
Cl$_2$ | Cl$^-$	2 Cl$^-$ = Cl$_2$(g, 1 atm) + 2 $\ominus$	− 1.3583
Pt | Ce^{+++}, Ce^{++++}	Ce^{+++} = Ce^{++++} + $\ominus$	− 1.55
Pt | Co^{++}, Co^{+++}	Co^{++} = Co^{+++} + $\ominus$	− 1.817

In every instance the sign of the emf refers to the *oxidation* reaction, and is positive when the oxidation is spontaneous and negative otherwise. Before using these $\mathcal{E}^0$ values in equation (29) it must first be ascertained whether the electrode reaction *as it actually occurs in the particular cell* is the oxidation or the reduction. For the oxidation the $\mathcal{E}^0$'s tabulated are

used as they are. However, for reduction the sign of each and every $\mathcal{E}^0$ given must be reversed. That is, for a process such as $K = K^+ + \ominus$ $\mathcal{E}^0_{25°}$ $= + 2.9241$, but for $K^+ + \ominus = K$ $\mathcal{E}^0_{25°} = - 2.9241$ volts. Similarly for $Cu = Cu^{++} + 2 \ominus$ $\mathcal{E}^0_{25°} = - 0.3441$ volt, while for $Cu^{++} + 2 \ominus = Cu$ $\mathcal{E}^0_{25°} = + 0.3441$ volt.

In equation (29) the factor $(2.3026\,RT)/\mathcal{F}$ is a constant for any temperature, and may readily be evaluated. To do this R must be taken in joules, namely $R = 8.315$ joules per degree. With this value of R and $\mathcal{F} = 96,500$ coulombs, this factor follows as

$$\frac{2.3026\,RT}{\mathcal{F}} = \frac{2.3026 \times 8.315\,T}{96,500}$$
$$= 1.9840 \times 10^{-4}\,T$$

Values of $1.9840^{-4}\,T$ for a number of convenient and frequently required temperatures are given in Table 4.

<div align="center">

TABLE 4

Values of $\dfrac{2.3026\,RT}{\mathcal{F}}$ at Various Temperatures

</div>

$t°$ C	$T°$ K	$\dfrac{2.3026\,RT}{\mathcal{F}} = 1.9840 \times 10^{-4}\,T$
0	273.15	0.05419
10	283.15	0.05618
15	288.15	0.05717
20	293.15	0.05816
25	298.15	0.05915
30	303.15	0.06014

Once the $\mathcal{E}^0$'s for various electrodes are available, they may readily be utilized to calculate the single electrode potentials. Specifically, suppose the emf of the zinc electrode

$$Zn \mid Zn^{++}(a = 0.1)$$
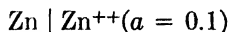
is sought at 25° C in a solution of zinc ions where the activity is 0.1. For *oxidation* the reaction at this electrode is

$$Zn(s) = Zn^{++}(a = 0.1) + 2 \ominus$$
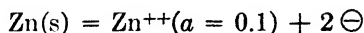
and hence on applying the Nernst equation we obtain for the single electrode potential, $\mathcal{E}_{Zn}$,

$$\mathcal{E}_{Zn} = \mathcal{E}^0_{Zn} - \frac{2.3026\,RT}{n\mathcal{F}} \log_{10} \frac{a_{Zn^{++}}}{a_{Zn}}$$

But for this reaction $n = 2$, $a_{Zn^{++}} = 0.1$, and $a_{Zn} = 1$, since zinc is a solid. Again, at 25° C, $(2.3026\,RT)/\mathcal{F} = 0.05915$, while $\mathcal{E}^0_{Zn}$ for oxida-

tion is $\mathcal{E}^0_{25^\circ} = +0.7611$ volt. Substituting these quantities into the expression for $\mathcal{E}_{Zn}$, we find that

$$\mathcal{E}_{Zn} = +0.7611 - \frac{0.05915}{2} \log_{10} \frac{0.1}{1}$$
$$= 0.7611 + 0.0296$$
$$= 0.7907 \text{ volt}$$

However, when the reaction at the electrode is a *reduction*, namely,

$$Zn^{++}(a = 0.1) + 2 \ominus = Zn(s)$$
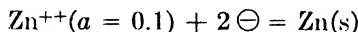
the Nernst equation yields for $\mathcal{E}_{Zn}$:

$$\mathcal{E}_{Zn} = \mathcal{E}^0_{Zn} - \frac{0.05915}{2} \log_{10} \frac{a_{Zn}}{a_{Zn^{++}}}$$
$$= \mathcal{E}^0_{Zn} - \frac{0.05915}{2} \log_{10} \frac{1}{0.1}$$

Further, for reduction $\mathcal{E}^0_{Zn} = -0.7611$, and therefore,

$$\mathcal{E}_{Zn} = -0.7611 - 0.0296$$
$$= -0.7907 \text{ volt}$$

i.e., the potential is the same as before but opposite in sign.

The Nernst Equation and Cell Emf's. Consider now a cell

$$Zn \mid Zn^{++}(a_{Zn^{++}}) \parallel Cl^-(a_{Cl^-}), Hg_2Cl_2(s) \mid Hg \qquad (30)$$
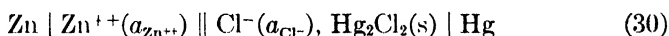
In this cell the zinc electrode is negative and undergoes an oxidation, while the calomel electrode is positive and suffers reduction. For the oxidation the electrode reaction is

$$Zn(s) = Zn^{++}(a_{Zn^{++}}) + 2 \ominus \qquad (31a)$$
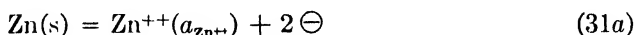
with an electrode emf given by the Nernst equation as

$$\mathcal{E}_{Zn} = \mathcal{E}^0_{Zn} - \frac{RT}{2\,\mathcal{F}} \ln a_{Zn^{++}} \qquad (31b)$$

Again, for the reduction at the calomel electrode, the reaction is

$$Hg_2Cl_2(s) + 2 \ominus = 2\,Hg(l) + 2\,Cl^-(a_{Cl^-}) \qquad (32a)$$
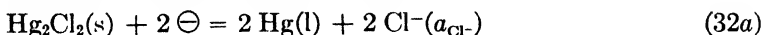
and the single electrode potential is

$$\mathcal{E}_C = \mathcal{E}^0_C - \frac{RT}{2\,\mathcal{F}} \ln a^2_{Cl^-} \qquad (32b)$$

On adding equations (31a) and (32a) we find the cell reaction to be

$$Zn(s) + Hg_2Cl_2(s) = 2\,Hg(l) + Zn^{++}(a_{Zn^{++}}) + 2\,Cl^-(a_{Cl^-}) \qquad (33a)$$
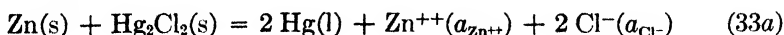
and similarly on adding equations (31b) and (32b) the cell emf $\mathcal{E}$ follows as

$$
\begin{aligned}
\mathcal{E} &= \mathcal{E}_{Zn} + \mathcal{E}_{C} \\
&= \left(\mathcal{E}_{Zn}^{0} - \frac{RT}{2\,\mathcal{F}} \ln a_{Zn^{++}} \right) + \left(\mathcal{E}_{C}^{0} - \frac{RT}{2\,\mathcal{F}} \ln a_{Cl^-}^2 \right) \\
&= (\mathcal{E}_{Zn}^{0} + \mathcal{E}_{C}^{0}) - \frac{RT}{2\,\mathcal{F}} \ln(a_{Zn^{++}} a_{Cl^-}^2)
\end{aligned}
\tag{33b}
$$

Instead of approaching the cell emf in this manner through the single electrode potentials, we may apply the Nernst equation (29) *directly* to the overall cell reaction as given in equation (33a). If this be done we find that the cell emf is

$$
\mathcal{E} = \mathcal{E}_{cell}^{0} - \frac{RT}{2\,\mathcal{F}} \ln(a_{Zn^{++}} a_{Cl^-}^2)
\tag{34}
$$

where $\mathcal{E}_{cell}^{0}$ is the *standard potential of the cell.* Comparison of equations (33b) and (34), which must be identical, immediately shows that $\mathcal{E}_{cell}^{0} = \mathcal{E}_{Zn}^{0} + \mathcal{E}_{C}^{0}$, i.e., that the $\mathcal{E}^0$ *value of the cell is the sum of the $\mathcal{E}^0$ values of the electrodes composing the cell.* This relation between the standard electrode potentials and the standard cell potential, namely,

$$
\mathcal{E}_{cell}^{0} = \mathcal{E}_{1}^{0} + \mathcal{E}_{2}^{0}
\tag{35}
$$

where $\mathcal{E}_{1}^{0}$ and $\mathcal{E}_{2}^{0}$ are the $\mathcal{E}^0$'s for the two electrodes, applies to any cell. Consequently, as soon as the standard emf's of the electrodes are known, that of the cell composed of these electrodes is also available through equation (35), and the Nernst equation may be applied directly to the overall reaction of the cell.

In employing equation (35) to obtain standard cell potentials from the normal electrode potentials given in Table 3, it should be remembered that each standard electrode potential must be prefixed by the *sign corresponding to the reaction as it occurs in the cell.* Thus the reaction at the zinc electrode in equation (30) is an oxidation, and therefore $\mathcal{E}_{Zn}^{0} = + 0.7611$ volt at 25° C. Again, the reaction at the calomel electrode is a reduction, and hence $\mathcal{E}_{C}^{0} = + 0.2680$ volt. Consequently, $\mathcal{E}_{cell}^{0}$ is

$$
\begin{aligned}
\mathcal{E}_{cell}^{0} &= 0.7611 + 0.2680 \\
&= 1.0291 \text{ volts}
\end{aligned}
$$

With this value of $\mathcal{E}_{cell}^{0}$ and $a_{Zn^{++}} = 0.1$, $a_{Cl^-} = 0.2$, the cell given in equation (30) would have, according to equation (34), the emf

$$
\begin{aligned}
\mathcal{E} &= \mathcal{E}_{cell}^{0} - \frac{RT}{2\,\mathcal{F}} \ln a_{Zn^{++}} a_{Cl^-}^2 \\
&= 1.0291 - \frac{0.05915}{2} \log_{10}(0.1)(0.2)^2 \\
&= 1.0291 + 0.0708 \\
&= 1.0999 \text{ volts}
\end{aligned}
$$

Standard Potentials and Equilibrium Constants. Besides their use for the calculation of single electrode and cell potentials, standard electrode and cell emf's may also be utilized for obtaining equilibrium constants. According to equation (28), ΔF^0 for any electrochemical process is given by

$$\Delta F^0 = -n\mathcal{F}\mathcal{E}^0 .$$

But we have seen in Chapter XII, equation (17), that ΔF^0 is also related to the equilibrium constant K_a of the process by the equation

$$\Delta F^0 = -RT \ln K_a$$

If these two expressions for ΔF^0 are equated, we obtain, for the relation between $\mathcal{E}^0$ and K_a,

$$-n\mathcal{F}\mathcal{E}^0 = -RT \ln K_a$$

and hence,
$$\mathcal{E}^0 = \frac{RT}{n\mathcal{F}} \ln K_a \tag{36}$$

With equation (36) K_a can readily be evaluated from $\mathcal{E}^0$; or, vice versa, when K_a is available, $\mathcal{E}^0$ may be calculated. Thus, for the reduction of stannous ions by thallium, $\mathcal{E}^0 = 0.1979$ volt at 25° C. For this process the equilibrium constant is, therefore,

$$0.1979 = \frac{0.05915}{1} \log_{10} K_a$$
$$\log_{10} K_a = \frac{0.1979}{0.05915} = 3.347$$
$$K_a = 2.22 \times 10^3$$

Since K_a for this reaction is

$$K_a = \frac{a_{Tl^+}}{a_{Sn^{++}}^{1/2}}$$

where the activities are those *at equilibrium*, this means that if thallium were added to a solution of stannous ions, the reaction would proceed and equilibrium would not be established until the activity of thallous ions becomes equal to

$$a_{Tl^+} = K_a \, a_{Sn^{++}}^{1/2}$$
$$= 2.22 \times 10^3 \, a_{Sn^{++}}^{1/2}$$

Once this condition is attained, the system will be in equilibrium, and no further chemical action will take place.

In this manner may be calculated the equilibrium conditions for any electrode or cell reaction from the appropriate $\mathcal{E}^0$ values. Qualitatively the results to be anticipated may be arrived at as follows. Inspection of

Table 3 shows that for the oxidation reactions listed above hydrogen $\mathcal{E}^0$ is positive, and hence for these the oxidation processes are spontaneous with respect to hydrogen. Further, the metals at the top of the table possess high values of $\mathcal{E}^0$, indicating that these substances are oxidized readily with a large decrease in free energy. Since any substance that is readily oxidized is a strong reducing agent, we may conclude that the metals high in the electromotive series of standard electrode potentials are powerful *reducing agents*, and that the reducing power of these decreases as we pass *down* the table. On the other hand, the substances below hydrogen in the table are not oxidized spontaneously, and for these, therefore, the reduction process is spontaneous. Furthermore, the reduction proceeds the more readily the lower the reaction is in the electromotive series. As substances that are easily reduced are strong oxidizing agents, we find thus at the bottom of the table the *powerful oxidants*, such as Ce^{++++}, Tl^{+++}, chlorine, bromine; and, as we pass *up* the table the oxidizing power decreases. Consequently extent of reaction will be most complete for chemical interaction between substances at the top and bottom of the table, and will decrease as the substances reacting lie closer and closer to each other. Thus, for reaction between lithium metal and ceric ions, namely,

$$Li(s) + Ce^{++++} = Ce^{+++} + Li^+$$
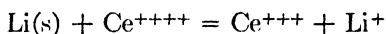
K_a at 25° C is 10^{77}, i.e., the reduction of ceric ions is complete, whereas, for the reaction between iron metal and cadmium ions,

$$Fe(s) + Cd^{++} = Fe^{++} + Cd(s)$$
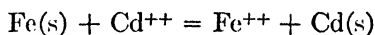
K_a is only 4.6, indicating that at equilibrium an appreciable quantity of cadmium ions is still present.

The relation of oxidizing and reducing power to the position of various substances in the electromotive series is of great help in analytical chemistry, for it allows a choice of oxidizing and reducing agents for various reactions. Theoretically any substance should be reduced by any other above it in the series, and oxidized by any one below it, provided the activities of all reactants are unity. As the latter condition is hardly ever encountered in practice, this expectation does not always materialize. In fact, the position of various substances near each other in the series may actually be shifted about by changes in the concentrations of the reactants. However, for substances considerably removed from each other in the series such a shift is not possible, and to such the above relation can be applied with effect.

Mechanism of Electrode Processes. In the preceding sections we have seen how potentials can be measured, and how by application of thermodynamic methods to these valuable information can be obtained

concerning the energy relations of the reactions occurring at the electrodes. However, these methods tell us nothing about the manner in which the emf arises, or where it is located; i.e., they tell us nothing about the intimate mechanism of the electrode or cell processes. Still, the question of the source of emf and the mechanism of its establishment is a subject of considerable interest, and warrants some consideration.

Since a single phase does not exhibit an emf, we may look for its source at the point of contact between any two phases. Thus, in any cell such as that shown in Fig. 6, consisting of a cadmium electrode in a solution of cadmium sulfate and a copper electrode in a solution of copper sulfate, the possible sources of the observed emf may be the cadmium-
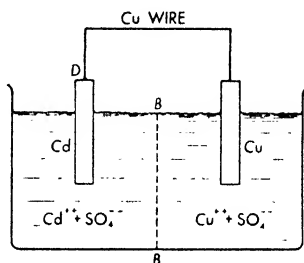
Fig. 6. Possible Seats of Emf in a Cell

cadmium sulfate and copper-copper sulfate interphases, the point of contact between the copper wire and the cadmium metal D, or the junction of the two solutions at B. Although, as will be shown, the junction of two dissimilar solutions does lead to a junction potential, the magnitude of the latter is usually relatively small compared to the total emf of the cell. Further, by appropriate experimental means it is possible either to minimize this potential to a point where it is negligible, or to construct cells of definite emf without liquid junctions. In either case junction potentials cannot possibly account for the observed emf's. We are left, therefore, with the metallic contact potentials or the interphase potentials between metals and their solutions, or both, as the possible sources of emf.

Originally Volta believed that the establishment of an emf between two electrodes was due to the contact potential between the metals composing the two electrodes. However, the fact that the potentials of electrodes depend on the concentration of the ions in which they are immersed led Walter Nernst to reject this view, and to propose in 1889 his *solution tension theory* of electrode processes. In this theory the seat of the emf is exclusively the interphase between electrodes and their solutions. According to Nernst, the emf established at any electrode is a result of three forces operating at the electrode. First there is the *solution tension* or *solution pressure* of the element which tends to drive the substance into solution, and as a result of which the element acquires a charge. Second, this tendency of the element to dissolve is opposed by the ions of the element already present in solution. A measure of this opposition is the *osmotic pressure* of the ions, which operates to deposit ions from the solution onto the element constituting the electrode.

Finally, there is an *electrostatic force* involved between the charged element and the ions in solution. This force always acts in such a direction as to bring about equilibrium among all the forces.

The manner in which the three forces postulated lead to the appearance of an emf at an electrode may be seen from the following examples.
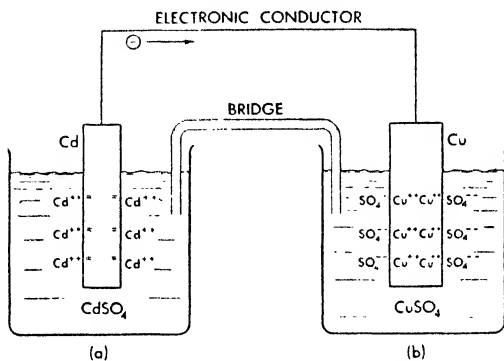
Fig. 7. Nernst Solution Tension Theory

Let us take first a cadmium electrode in a solution of cadmium ions, Fig. 7(a). Because of the solution tension of this metal, cadmium tends to dissolve as cadmium ions. To do this each cadmium atom must ionize on the metal surface, leave two electrons behind, and then pass into solution. As a consequence of this operation the metal acquires a negative charge, while the positive ions entering the solution cause it to become positively charged in the immediate vicinity of the electrode. Were this process to continue indefinitely, the metal would eventually dissolve. However, this cannot happen, for the entrance of the ions into the solution is opposed by the osmotic pressure of the cadmium ions already present. Again, as the cadmium ions enter the solution they cannot move any appreciable distance from the electrode because they are attracted by the negative charges on the electrode. The result is that they arrange themselves next to the electrode, as shown in the figure, to produce a *Helmholtz double layer* of charges, with the negative electrons on the metal and the positive ions in solution. This array of charges produces an electrostatic force which opposes the passage of positive ions through it. Hence this force acts with the osmotic pressure against the solution tension. When this electrostatic force becomes sufficiently large, so that the sum of it and the osmotic pressure becomes equal to the solution tension of the metal, all action ceases, and the electrode reaches a state of equilibrium with respect to the solution. But in attaining this equilibrium state the electrode has acquired a negative charge which is opposite to that of the solution, and hence

there must exist a potential difference across the metal-solution interphase. This potential difference Nernst considers to be the equilibrium emf of the electrode.

This state of affairs obtains when the solution tension of the metal is greater than the osmotic pressure of the solution of its ions. When the reverse is true, as in the copper-copper sulfate electrode, then we have the situation shown in Fig. 7(b). Since here the osmotic pressure of the solution is greater than the solution tension of the metal, Cu^{++} ions, instead of passing into solution from the metal, are forced to deposit on the surface of the copper. Because of these ions on the metal, the electrode acquires a positive charge which attracts toward itself the excess negative ions in solution, namely, SO_4^{--}, to form again a Helmholtz double layer. However, the arrangement of the charges in the double layer is now opposite to that in the cadmium; i.e., it is such as to favor the passage of copper ions from the metal to the solution and to oppose their passage in the opposite direction. In other words, the electrostatic force of the double layer operates with the weaker solution tension of the metal against the stronger osmotic pressure of the solution. When the sum of the forces of electrostatic attraction and solution tension becomes equal to that of the osmotic pressure, the electrode attains its equilibrium, and the potential difference between the metal and its solution at this stage becomes the emf of the electrode.

If the two solutions are brought together now, and the cadmium electrode with its excess of electrons is connected by a metallic conductor with the copper electrode, where a deficiency of electrons exists, electrons will flow through the external circuit from the cadmium to the copper. By removing electrons from the cadmium electrode the equilibrium is disturbed, and hence more cadmium ions will pass into solution to reestablish the equilibrium. Again, by introducing electrons to the copper electrode copper ions will be removed to form copper atoms, and to readjust the electrode to its equilibrium more copper ions will have to pass from the solution to the electrode. The net result is that by joining the two electrodes we obtain electrical energy at the expense of the two electrode reactions,

$$Cd(s) = Cd^{++} + 2 \ominus$$
$$Cu^{++} + 2 \ominus = Cu(s)$$
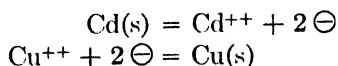
or the overall cell reaction

$$Cd(s) + Cu^{++} = Cd^{++} + Cu(s)$$
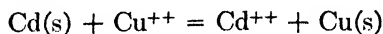
These ideas may be utilized to explain the operation of other electrodes and other cells. Moreover, by applying thermodynamic arguments to the theory outlined, Nernst was able to deduce an equation identical in

form with equation (29) for the potential as a function of concentration. However, application of quantum mechanical considerations [1] to electrode processes seems to indicate that Volta potentials between metals are not negligible, and that they enter into the emf's of cells in a very intimate manner. According to Gurney, the Volta potentials are the *cause* which leads to the Nernst mechanism as the effect. The arguments for this statement are beyond our scope. Suffice it to say that these newer ideas show that, although cell mechanisms may involve both the Volta and Nernst potentials, the experimentally observed emf's of reversible cells arise ultimately from the chemical reactions producing them, and measure the free energy changes of these.

Classification of Electrodes. In electrochemical work the cells encountered involve various electrodes depending on the purpose at hand. These various electrodes may. be grouped into seven types, namely,

(1) Metal-metal ion electrodes.
(2) Amalgam electrodes.
(3) Nonmetal nongas electrodes.
(4) Gas electrodes.
(5) Metal-insoluble salt electrodes.
(6) Metal-insoluble oxide electrodes.
(7) Oxidation-reduction electrodes.

We proceed now to a discussion of each of these.

Metal-Metal Ion Electrodes. Electrodes of this type involve a metal in equilibrium with a solution of its ions. They have already been described in some detail. Examples are the zinc, copper, cadmium, and sodium electrodes. All these electrodes operate on the general reaction

$$M = M^{+n} + n \ominus \tag{37}$$

for which the electrode potential equation is given by

$$\mathcal{E}_M = \mathcal{E}_M^0 - \frac{RT}{n\mathcal{F}} \ln a_{M^{+n}} \tag{38}$$

Each of these electrodes is said to be *reversible* to its own ions; i.e., the potential of each of these electrodes is sensitive to and is determined by the activities of its own ions in the solution in which the metal is immersed. Thus, the zinc electrode is reversible to zinc ions, the iron electrode to ferrous ions, the tin electrode to stannous ions, etc.

Amalgam Electrodes. It is quite common practice to substitute for the pure metals in metal-metal ion electrodes solutions of the metal

---

[1] Gurney, "Ions in Solution," Cambridge University Press, Cambridge, Eng., 1936. For simple discussion see Dole, "Experimental and Theoretical Electrochemistry," McGraw-Hill Book Company, Inc., New York, 1935, Chapter XXIX.

in mercury, namely, *amalgams*. Amalgams of metals more active than mercury behave essentially as do the pure metals, the only difference being that the activity of the metal is lowered somewhat by dilution by the mercury. Electrodes of these amalgams are preferred frequently because equilibrium can as a rule be established much more rapidly with them than with the pure metals, and because they are more readily reversible. Again, with metals such as sodium, potassium, or calcium the activity in aqueous solutions is too great for direct use. However, by converting these metals to amalgams the activity can be moderated sufficiently to permit measurements in presence of water. Still another factor in favor of the amalgams is that small quantities of impurities which may cause erratic behavior in pure metals are often diluted enough by amalgamation to yield satisfactory and reproducible results.

As an example of electrodes of this class may be taken the lead amalgam electrode, consisting of the lead amalgam, written Pb(Hg), dipping into a solution of plumbous ions, namely, Pb(Hg) | Pb^{++}($a_{Pb^{++}}$). For this electrode the reaction is

$$Pb(Hg) = Pb^{++}(a_{Pb^{++}}) + 2 \ominus \tag{39}$$

and hence the electrode potential $\mathcal{E}_a$ is given by

$$\mathcal{E}_a = \mathcal{E}_{Pb}^0 - \frac{RT}{2\,\mathfrak{F}} \ln \frac{a_{Pb^{++}}}{a_{Pb}} \tag{40}$$

where $\mathcal{E}_{Pb}^0$ is the standard electrode potential of the pure lead electrode, $a_{Pb^{++}}$ is the activity of plumbous ions in solution, and a_{Pb} is the activity of *metallic lead in the amalgam*. In general a_{Pb} is not unity. The usual procedure is to write equation (40) as

$$\mathcal{E}_a = \left(\mathcal{E}_{Pb}^0 + \frac{RT}{2\,\mathfrak{F}} \ln a_{Pb}\right) - \frac{RT}{2\,\mathfrak{F}} \ln a_{Pb^{++}}$$

$$= \mathcal{E}_a^0 - \frac{RT}{2\,\mathfrak{F}} \ln a_{Pb^{++}} \tag{41}$$

where $\mathcal{E}_a^0$ is the standard potential of the given amalgam, and to evaluate $\mathcal{E}_a^0$ first. Then, in order to transform $\mathcal{E}_a^0$ to $\mathcal{E}_{Pb}^0$, the emf of the amalgam is measured against pure lead when both are immersed in a solution of plumbous ions of the *same* concentration. Since for pure lead the single electrode potential equation is

$$\mathcal{E}_{Pb} = \mathcal{E}_{Pb}^0 - \frac{RT}{2\,\mathfrak{F}} \ln a_{Pb^{++}}$$

while for the amalgam it is equation (41), the difference in potential $\mathcal{E}$ between the lead and amalgam electrodes follows as

$$
\begin{aligned}
\mathcal{E} &= \mathcal{E}_{\text{Pb}} - \mathcal{E}_a \\
&= \left(\mathcal{E}_{\text{Pb}}^0 - \frac{RT}{2\,\mathcal{F}} \ln a_{\text{Pb}^{++}} \right) - \left(\mathcal{E}_a^0 - \frac{RT}{2\,\mathcal{F}} \ln a_{\text{Pb}^{++}} \right) \\
&= \mathcal{E}_{\text{Pb}}^0 - \mathcal{E}_a^0
\end{aligned}
\tag{42}
$$

Consequently, with $\mathcal{E}$ and $\mathcal{E}_a^0$ known $\mathcal{E}_{\text{Pb}}^0$ may be evaluated, and the emf data obtained with the amalgam may be reduced to the pure metal. Thus Carmody [1] found that for the particular lead amalgam he employed in his studies the standard electrode potential of $\text{Pb}(\text{Hg}) \mid \text{Pb}^{++}$ was 0.1207 volt at 25° C, while the emf of a cell composed of this amalgam and pure lead in a solution of plumbous ions was 0.0058 volt. The standard potential of the $\text{Pb} \mid \text{Pb}^{++}$ electrode at 25° C, $\mathcal{E}_{\text{Pb}}^0$, is then, according to equation (42),

$$
\begin{aligned}
\mathcal{E}_{\text{Pb}}^0 &= \mathcal{E} + \mathcal{E}_a^0 \\
&= 0.0058 + 0.1207 \\
&= 0.1265 \text{ volt}
\end{aligned}
$$

Nonmetal Nongas Electrodes. Electrodes may be constructed from nonmetals as well as from metals. For example, solid iodine in presence of iodide ions will serve as an electrode reversible to iodide ions. Similarly liquid bromine in presence of bromide ions will act as an electrode reversible to bromide ions. In assembling such electrodes, the condensed nonmetal phase is placed in the bottom of a glass vessel, and over it is poured a solution containing the ions to which the substance is reversible. Electric contact is made by means of an inert metal, such as platinum, sealed into the bottom of the vessel or inserted through the top so as to touch the condensed phase.

The electrode reaction for the iodine electrode is

$$
\frac{1}{2} \text{I}_2(\text{s}) + \ominus = \text{I}^-(a_{\text{I}^-})
\tag{43}
$$

with the electrode emf given by

$$
\mathcal{E}_{\text{I}_2} = \mathcal{E}_{\text{I}_2}^0 - \frac{RT}{\mathcal{F}} \ln a_{\text{I}^-}
\tag{44}
$$

In a like manner the single electrode process of the bromine electrode is

$$
\frac{1}{2} \text{Br}_2(\text{l}) + \ominus = \text{Br}^-(a_{\text{Br}^-})
\tag{45}
$$
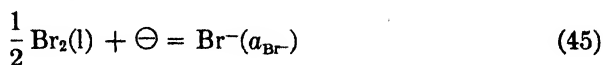
with the emf given by

$$
\mathcal{E}_{\text{Br}_2} = \mathcal{E}_{\text{Br}_2}^0 - \frac{RT}{\mathcal{F}} \ln a_{\text{Br}^-}
\tag{46}
$$

[1] Carmody, J. Amer. Chem. Soc., **51**, 2905 (1929).

For both these electrodes the *reduction* is spontaneous, and for both, therefore, $\mathcal{E}^0$ is positive for the reactions given in equations (43) and (45).

Gas Electrodes. Gas electrodes consist of a gas bubbling about an inert metal wire or foil immersed in a solution containing ions to which the gas is reversible. The function of the metal wire or foil, which usually is platinized platinum, is to facilitate establishment of equilibrium between the gas and its ions and to serve as the electric contact for the electrode.

Among the gas electrodes are the hydrogen electrode, which is reversible to hydrogen ions, the chlorine electrode, reversible to chloride ions, and the oxygen electrode, whose emf depends on the activity of hydroxyl ions. However, though the first two electrodes can be made reversible, no suitable electrode material has so far been found which can catalyze satisfactorily the establishment of equilibrium between oxygen and hydroxyl ions. Whatever precise information is available on the latter electrode has been obtained, not by direct emf measurements, but by calculation from suitable standard free energy data obtained from various other sources.

The fundamental electrode reaction for the hydrogen electrode is

$$\frac{1}{2}\,H_2(g,\,P_{H_2}) = H^+(a_{H^+}) + \ominus \tag{47}$$
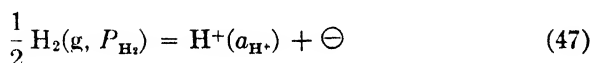
Since the activity of hydrogen gas is at relatively low pressures equal to the pressure P_{H_2} of the gas in atmospheres, and since the activity of hydrogen ions a_{H^+} depends on the concentration of the latter in solution, the single electrode potential of the hydrogen electrode must be determined both by the pressure of the hydrogen about the electrode, and the activity of hydrogen ions in solution. Namely,

$$\mathcal{E}_{H_2} = \mathcal{E}^0_{H_2} - \frac{RT}{\mathcal{F}} \ln \frac{a_{H^+}}{P_{H_2}^{1/2}} \tag{48}$$

But $\mathcal{E}^0_{H_2}$, i.e., the *emf of the hydrogen electrode at 1 atmosphere hydrogen pressure and at unit activity of hydrogen ions*, is the reference of all emf measurements, and is taken by definition to be zero at all temperatures. Consequently equation (48) becomes

$$\begin{aligned}
\mathcal{E}_{H_2} &= -\frac{RT}{\mathcal{F}} \ln \frac{a_{H^+}}{P_{H_2}^{1/2}} \\
&= -\frac{RT}{\mathcal{F}} \ln a_{H^+} + \frac{RT}{\mathcal{F}} \ln P_{H_2}^{1/2}
\end{aligned} \tag{49}$$

Further, when the pressure of hydrogen is 1 atmosphere, $P_{H_2}^{1/2} = 1$, and $\ln P_{H_2}^{1/2} = 0$. Then,

$$\mathcal{E}_{H_2} = -\frac{RT}{\mathcal{F}} \ln a_{H^+} \tag{50}$$

and the *hydrogen electrode becomes strictly dependent only on the activity of hydrogen ions in solution,* i.e., the pH. This fact is of great importance, for it permits the estimation of pH from emf measurements. Reference in greater detail to this application of the hydrogen electrode will be made later.

The chlorine electrode behaves analogously to the hydrogen electrode in that its emf is determined both by the pressure of chlorine gas about the electrode and the activity of chloride ions in the solution with which it is in contact. For this electrode the spontaneous reaction is the reduction,

$$\frac{1}{2}\,Cl_2(g,\,P_{Cl_2}) + \ominus = Cl^-(a_{Cl^-}) \tag{51}$$
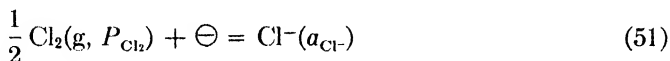
which yields for the electrode equation

$$\mathcal{E}_{Cl_2} = \mathcal{E}^0_{Cl_2} - \frac{RT}{\mathcal{F}} \ln \frac{a_{Cl^-}}{P^{1/2}_{Cl_2}} \tag{52}$$

Unlike $\mathcal{E}^0_{H_2}$, $\mathcal{E}^0_{Cl_2}$ is not zero, but is in fact equal to 1.3583 volts at 25° C.

Metal-Insoluble Salt Electrodes. Electrodes of this type are extremely important in electrochemistry, and are encountered very frequently. In this category fall the various calomel electrodes, the silver-silver chloride electrode, the lead-lead sulfate electrode, the silver-silver bromide electrode, and others.

The common characteristic of electrodes in this class is that they all consist of a metal in contact with one of its difficultly soluble salts, and a solution containing the ion present in the salt *other than the metal.* For instance, a mercury-mercurous chloride electrode (calomel) consists of mercury, solid mercurous chloride, and a solution containing chloride ions, such as potassium chloride, sodium chloride, etc. Similarly a silver-silver chloride electrode is composed of a silver wire coated with silver chloride and immersed in a solution of chloride ions. Again, the lead-lead sulfate electrode may be either pure lead or a lead amalgam covered with crystals of lead sulfate, and surrounded by solutions of sulfates, such as sodium sulfate, sulfuric acid, or magnesium sulfate. In all instances *these electrodes are reversible to the ions other than those of the metal present in the insoluble salt.* That is, the calomel and silver-silver chloride electrodes are reversible to chloride ions, the lead-lead sulfate electrode to sulfate ions, the silver-silver bromide electrode to bromide ions, etc. The reason may be seen from the following discussion, in which we shall take the silver-silver chloride electrode as a typical member of this class.

Consider a silver electrode dipping into a solution possessing chloride ions and saturated with silver chloride, i.e., solid silver chloride is present. If we imagine this electrode to act as an ordinary metal-metal ion

electrode, silver ions will pass from the electrode into the solution according to the reaction

$$Ag(s) = Ag^+ + \ominus \tag{a}$$

However, since the solution is saturated with silver chloride, the introduction of any silver ions will upset the requirements of the solubility product principle for this salt. Consequently, in order to reestablish the solubility equilibrium, silver ions will have to combine with chloride ions to precipitate solid silver chloride according to the relation

$$Ag^+ + Cl^- = AgCl(s) \tag{b}$$

Moreover, as the overall electrode reaction is the sum of all the processes occurring at the electrode, the overall electrode reaction will have to be the sum of (a) and (b), or

$$Ag(s) + Cl^- = AgCl(s) + \ominus \tag{53}$$

For this reaction the electrode potential equation is

$$\mathcal{E}_{Ag-AgCl} = \mathcal{E}^0_{Ag-AgCl} - \frac{RT}{\mathcal{F}} \ln \frac{1}{a_{Cl^-}} \tag{54}$$

and hence this electrode is reversible to *chloride ions*, i.e., its emf is controlled by the activity of chloride ions in the solution.

In an analogous manner may be deduced the electrode reactions for the other electrodes of this type. Thus, we find that for the calomel electrode the reaction is
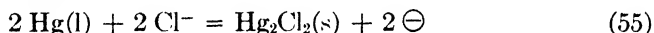
$$2\,Hg(l) + 2\,Cl^- = Hg_2Cl_2(s) + 2\ominus \tag{55}$$

for the lead-lead sulfate electrode,

$$Pb(s) + SO_4^{--} = PbSO_4(s) + 2\ominus \tag{56}$$

for the silver-silver bromide electrode,

$$Ag(s) + Br^- = AgBr(s) + \ominus \tag{57}$$
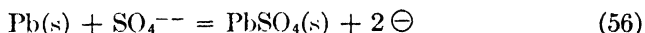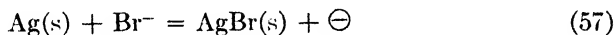
and so on. Reference to Table 3 will reveal that for the reactions given in equations (53), (55), and (57) the spontaneous processes are not the oxidations but the reductions; namely, the spontaneous reactions are the reverse of those shown. On the other hand, in the silver-silver iodide and lead-lead sulfate electrodes the oxidations proceed spontaneously.

Metal-Insoluble Oxide Electrodes. Metal-insoluble oxide electrodes are very similar to the metal-insoluble salt type, except that the difficultly soluble salt in the latter case is replaced by a difficultly soluble oxide. Such electrodes are reversible to *either hydrogen or hydroxyl ions.* The only important electrode in this category is the antimony-antimony

trioxide electrode, consisting of a stick of antimony covered with a thin layer of the oxide resulting from oxidation of the surface by oxygen.

The operation of the antimony electrode may be explained as follows. When a stick of the metal is placed in a solution containing hydrogen, and hence also hydroxyl, ions, antimony dissolves to yield Sb^{+++}, namely,

$$Sb(s) = Sb^{+++} + 3 \ominus$$
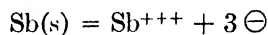
However, since the oxide is insoluble, Sb^{+++} and OH^- ions will combine to form $Sb_2O_3(s)$ according to the reaction

$$2\,Sb^{+++} + 6\,OH^- = Sb_2O_3(s) + 3\,H_2O(l).$$
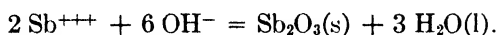
On doubling now the first of these equations and adding to the second, we obtain, for the overall electrode reaction,

$$2\,Sb(s) + 6\,OH^- = Sb_2O_3(s) + 3\,H_2O(l) + 6 \ominus \qquad (58)$$
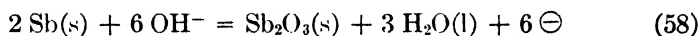
for which the single electrode emf is given by

$$\mathcal{E}_{Sb} = \mathcal{E}^0_{Sb} - \frac{RT}{6\,\mathcal{F}} \ln \frac{1}{a^6_{OH^-}}$$
$$= \mathcal{E}^0_{Sb} - \frac{RT}{\mathcal{F}} \ln \frac{1}{a_{OH^-}} \qquad (59)$$

In other words, the antimony-antimony trioxide electrode is reversible to *hydroxyl ions*. To show how this electrode is reversible to *hydrogen ions* as well, we may introduce the equilibrium

$$6\,H_2O(l) = 6\,H^+ + 6\,OH^-$$
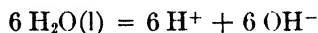
On combining the latter with equation (58) we obtain for the electrode reaction

$$2\,Sb(s) + 3\,H_2O(l) = Sb_2O_3(s) + 6\,H^+ + 6 \ominus \qquad (60)$$
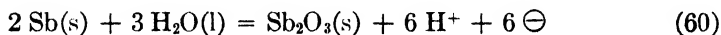
and hence $\mathcal{E}_{Sb}$ is also given by

$$\mathcal{E}_{Sb} = \mathcal{E}^{0\prime}_{Sb} - \frac{RT}{6\,\mathcal{F}} \ln a^6_{H^+}$$
$$= \mathcal{E}^{0\prime}_{Sb} - \frac{RT}{\mathcal{F}} \ln a_{H^+} \qquad (61)$$

$\mathcal{E}^0_{Sb}$ and $\mathcal{E}^{0\prime}_{Sb}$ are not the same numerically, but differ by a constant which can readily be shown to be $(-\,RT/\mathcal{F}) \ln K_w$, where K_w is the ion product of water; i.e.,

$$\mathcal{E}^0_{Sb} = \mathcal{E}^{0\prime}_{Sb} - \frac{RT}{\mathcal{F}} \ln K_w \qquad (62).$$

For equation (61), $\mathcal{E}^0_{Sb}$ at 25° C has been found to be -0.1445 volt. Consequently, $\mathcal{E}^0_{Sb}$ for equation (59) follows from equation (62) as

$$\mathcal{E}^0_{Sb} = -0.1445 - .05915 \log_{10} (1.008 \times 10^{-14})$$
$$= -0.1445 + 0.8279$$
$$= +0.6834 \text{ volt}$$

Oxidation-Reduction Electrodes. Although every electrode reaction involves an oxidation or a reduction, the term *oxidation-reduction electrodes* is used to designate electrodes in which the emf results from the presence of *ions* of a substance in two different stages of oxidation. When a platinum wire is inserted into a solution containing both ferrous and ferric ions it is found that the wire acquires a potential. The same is true of solutions of cerous-ceric ions, stannous-stannic ions, manganous-permanganate ions, etc. These electrode emf's arise from the tendency of ions in one state of oxidation to pass over into a second more stable state. The function of the platinum wire is merely to "pick up" the potential corresponding to this tendency toward a free energy decrease, and to serve as the electrical contact of the electrode.

The general electrode reaction for all electrodes of the oxidation-reduction type may be written as

$$A^{n_1}(a_1) + n \ominus = A^{n_2}(a_2) \tag{63}$$

where n_1 is the valence in the higher stage of oxidation, n_2 the valence in the lower stage, while $n = n_1 - n_2$ is the *valence change* attending the electrode process. From equation (63) the general emf equation follows as

$$\mathcal{E} = \mathcal{E}^0 - \frac{RT}{n\mathcal{F}} \ln \frac{a_2}{a_1} \tag{64}$$

i.e., the emf of oxidation-reduction electrodes depends only on the *ratio of the activities* of the two ions, and not at all on the absolute magnitudes of these. Thus, for the ferric-ferrous electrode, designated symbolically as $Pt \mid Fe^{+++}, Fe^{++}$, the electrode reaction is

$$Fe^{+++}(a_{Fe^{+++}}) + \ominus = Fe^{++}(a_{Fe^{++}}) \tag{65}$$
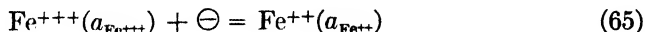
and the electrode emf is given by

$$\mathcal{E} = \mathcal{E}^0 - \frac{RT}{\mathcal{F}} \ln \frac{a_{Fe^{++}}}{a_{Fe^{+++}}} \tag{66}$$

In Table 3 are given $\mathcal{E}^0$ values for several electrodes of this type at 25° C.

Electrochemical Cells. When various electrodes are combined to form cells, two general classes of cells may result, namely: (1) *chemical*

cells, in which the emf is due to a *chemical reaction* occurring within the cell, and (2) *concentration cells*, in which the emf is due to the free energy decrease attending the *transfer of matter* from one part of the cell to another. Further, each of these types of cells may or may not involve a liquid junction; or, as it is usually said, the cell may or may not have *transference*. On this basis we arrive at the following classification of electrochemical cells:

Electrochemical Cells

Chemical Cells   Concentration Cells

| With | Without | With | Without |
| Transference | Transference | Transference | Transference |
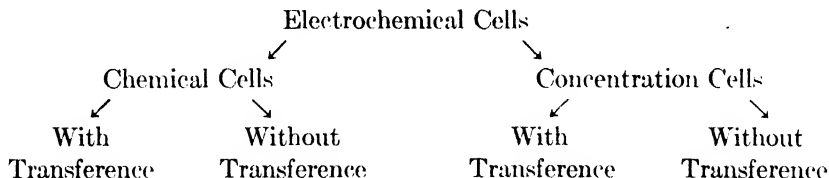
Chemical Cells without Transference. To construct a chemical cell which is to have no liquid junction or transference, two electrodes and an electrolyte must be selected such that *one of the electrodes is reversible to the cation, the other to the anion of the electrolyte*. For instance, if we desire a cell without transference employing hydrochloric acid as the electrolyte, we must use an electrode reversible to hydrogen ions, namely, the hydrogen electrode, and an electrode reversible to chloride ions. The latter may be either a silver-silver chloride, a mercury-mercurous chloride, or chlorine electrode, all of which are reversible to this ion. Again, for sulfuric acid as the electrolyte the electrodes will have to be hydrogen and either lead-lead sulfate or mercury-mercurous sulfate which are reversible to sulfate ions. Finally, for an electrolyte such as zinc bromide we shall have to employ zinc as one of the electrodes, and either the bromine, the mercury-mercurous bromide, or silver-silver bromide as the electrode for the bromide ions. Where a choice of electrodes is possible, as in the latter instance, the electrode that can be handled most conveniently is selected. Although the particular electrode chosen determines the overall cell reaction, and hence the $\mathcal{E}^0$ value of the cell, it is characteristic of all such possibilities that they give the same dependence of the emf of the cell on the activity of the electrolyte.

As a typical example of chemical cells without transference we shall take the cell

$$H_2(g, P_{H_2}) \mid HCl(a_{HCl}), AgCl(s) \mid Ag \qquad (67)$$

consisting of the hydrogen and silver-silver chloride electrodes in hydrochloric acid as the electrolyte. Since both these electrodes are reversible to the two ions of the electrolyte, they may be immersed directly in the acid to yield a cell possessing no liquid junction. The entire emf of the cell is composed then only of the emf's existing at the electrode-solution interphases.

Presupposing that cell (67) is written correctly, i.e., that the hydrogen electrode is negative, as it actually is, we obtain, for the oxidation at this electrode,

$$\frac{1}{2} H_2(g, P_{H_2}) = H^+(a_{H^+}) + \ominus \tag{68a}$$
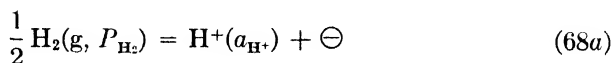
and
$$\mathcal{E}_{H_2} = -\frac{RT}{\mathcal{F}} \ln \frac{a_{H^+}}{P_{H_2}^{1/2}} \tag{68b}$$

Again, for the reduction at the silver-silver chloride electrode we have

$$AgCl(s) + \ominus = Ag(s) + Cl^-(a_{Cl^-}) \tag{69a}$$
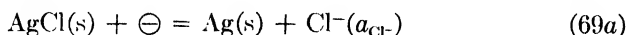
and therefore,
$$\mathcal{E}_{Ag-AgCl} = \mathcal{E}^0_{Ag-AgCl} - \frac{RT}{\mathcal{F}} \ln a_{Cl^-} \tag{69b}$$

Adding now equations (68a) and (69a), the cell reaction follows as

$$\frac{1}{2} H_2(g, P_{H_2}) + AgCl(s) = Ag(s) + H^+(a_{H^+}) + Cl^-(a_{Cl}) \tag{70a}$$
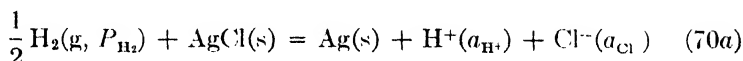
while the cell emf on combination of equations (68b) and (69b) is

$$
\begin{aligned}
\mathcal{E}_{cell} &= \mathcal{E}^0_{Ag-AgCl} - \frac{RT}{\mathcal{F}} \ln \frac{(a_{H^+} \cdot a_{Cl^-})}{P_{H_2}^{1/2}} \\
&= \mathcal{E}^0_{Ag-AgCl} - \frac{RT}{\mathcal{F}} \ln \frac{a_{HCl}}{P_{H_2}^{1/2}} \tag{70b}
\end{aligned}
$$

In the last equation a_{HCl}, the activity of the electrolyte as a whole, has been substituted for its equivalent [1] $a_{H^+} a_{Cl^-}$.

Equation (70a) indicates that the emf of the cell given in equation (67) results from a chemical reaction, namely, the reduction of silver chloride by hydrogen gas to form solid silver and hydrochloric acid ($H^+ + Cl^-$) in *solution*. Since this cell possesses as well no liquid junction, it constitutes a chemical cell without transference. Further, equation (70b) shows that the emf of this cell depends on the activity of the acid in solution and on the pressure of hydrogen gas. When the latter is 1 atmosphere, as it usually is under experimental conditions, equation (70b) reduces to

$$\mathcal{E}_{cell} = \mathcal{E}^0_{AgCl} - \frac{RT}{\mathcal{F}} \ln a_{HCl} \tag{71}$$

and the emf becomes dependent only on the activity of the hydrochloric acid in solution.

[1] See Chapter XVI, equation (2).

By following the steps outlined above the student can readily verify for himself that such cells as

$$H_2(g, P_{H_2}) \mid H_2SO_4(a_{H_2SO_4}), Hg_2SO_4(s) \mid Hg \tag{72a}$$

$$Cd \mid CdSO_4(a_{CdSO_4}), Hg_2SO_4(s) \mid Hg \tag{73a}$$
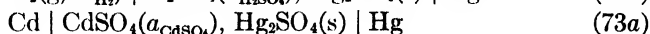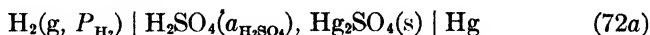
are chemical cells without transference for which the cell reactions are respectively:

$$H_2(g, P_{H_2}) + Hg_2SO_4(s) = 2\,Hg(l) + H_2SO_4(a_{H_2SO_4}) \tag{72b}$$

$$Cd(s) + Hg_2SO_4(s) = 2\,Hg(l) + CdSO_4(a_{CdSO_4}) \tag{73b}$$
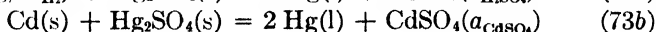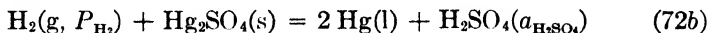
Finally, the emf equation for (72b) at 1 atmosphere hydrogen pressure follows as

$$\mathcal{E} = \mathcal{E}^0_{Hg-Hg_2SO_4} - \frac{RT}{2\,\mathcal{F}} \ln a_{H_2SO_4} \tag{72c}$$

while for (73b) $\mathcal{E}$ is

$$\mathcal{E} = \mathcal{E}^0_{cell} - \frac{RT'}{2\,\mathcal{F}} \ln a_{CdSO_4} \tag{73c}$$

where $\mathcal{E}^0_{cell}$ is the sum of the standard electrode potentials of the cadmium and mercury-mercurous sulfate electrodes.

Application of Chemical Cells without Transference. Chemical cells without transference find extensive application in electrochemistry for the evaluation of standard electrode potentials of cells and electrodes, and for the determination from emf data of the activity coefficients of various electrolytes. To illustrate the method by which these quantities are arrived at from the experimental emf's of such cells, let us take specifically the cell represented in equation (67), for which the cell emf at 1 atmosphere hydrogen pressure is given by equation (71). The problem to solve in this instance is: How can $\mathcal{E}^0_{Ag-AgCl}$ and the activity coefficients of hydrochloric acid solutions of various molalities be evaluated from a series of emf measurements made on cells such as (67) with different concentrations of the acid?

First, according to equation (15), Chapter XVI, the activity of the binary uni-univalent electrolyte hydrochloric acid is related at any molality m to the mean activity coefficient γ by the expression

$$a_{HCl} = m^2\gamma^2$$

Substituting this value of a_{HCl} into equation (71) and rearranging terms, we obtain

$$\mathcal{E}_{cell} = \mathcal{E}^0_{Ag-AgCl} - \frac{RT}{\mathcal{F}} \ln (m^2\gamma^2)$$

$$= \mathcal{E}^0_{Ag-AgCl} - \frac{2\,RT}{\mathcal{F}} \ln m - \frac{2\,RT}{\mathcal{F}} \ln \gamma$$

$$\left(\mathcal{E}_{cell} + \frac{2\,RT}{\mathcal{F}} \ln m \right) = \mathcal{E}^0_{Ag-AgCl} - \frac{2\,RT}{\mathcal{F}} \ln \gamma \tag{74}$$

All the quantities on the left-hand side of equation (74) are experimentally available, and hence γ could be calculated from $\mathcal{E}_{cell}$ and m as soon as $\mathcal{E}^0_{Ag-AgCl}$ is known. To determine the latter quantity $(\mathcal{E}_{cell} + 2RT/\mathcal{F} \ln m)$ is plotted against $\sqrt{\mu}$, i.e., the square root of the ionic strength of the solution, which in this case is identical with $\sqrt{m}$, and the plot is extrapolated to $\sqrt{\mu} = \sqrt{m} = 0$. To make this extrapolation with certainty, emf data for dilute solutions are necessary. As at $m = 0$, $\gamma = 1$, and $\ln \gamma = 0$, the last term on the right in equation (74) becomes thus zero, and the value of the extrapolated ordinate yields immediately $\mathcal{E}^0_{Ag-AgCl}$. Once this constant is known, the difference between $(\mathcal{E}_{cell} + 2RT/\mathcal{F} \ln m)$ and $\mathcal{E}^0_{Ag-AgCl}$ at each value of m gives immediately $-2RT/\mathcal{F} \ln \gamma$, and hence γ.

In Table 5 column 1 gives the molalities of a number of relatively dilute solutions of hydrochloric acid, column 2 the emf's at 25° C obtained experimentally for these with cells of the type under discussion. By calculating from these data the quantity $(\mathcal{E}_{cell} + 2RT/\mathcal{F} \ln m) = (\mathcal{E}_{cell} + 0.11830 \log_{10} m)$, and plotting it against $\sqrt{m}$, it is found that on extrapolation to $\sqrt{m} = 0$ the ordinate is 0.2225 volt. Consequently, $\mathcal{E}^0_{Ag-AgCl}$ at 25° C is 0.2225 volt. Utilizing this value of $\mathcal{E}^0_{Ag-AgCl}$ and the various values of $\mathcal{E}_{cell}$ and m in equation (74), the activity coefficients of hydrochloric acid are found to be those listed in column 3 of Table 5.

<div align="center">TABLE 5</div>

<div align="center">ACTIVITY COEFFICIENTS OF HCl FROM EMF'S OF THE CELL
$H_2 \mid HCl, AgCl(s) \mid Ag$ AT 25° C [1]</div>

m_{HCl}	$\mathcal{E}_{cell}$ (Volts)	Activity Coefficient γ
0.003215	0.52053	0.942
0.004488	0.50384	0.933
0.005619	0.49257	0.926
0.007311	0.47948	0.917
0.009138	0.46860	0.909
0.011195	0.45861	0.903
0.013407	0.44974	0.895
0.01710	0.43783	0.884
0.02563	0.41824	0.866
0.05391	0.38222	0.829
0.1238	0.34199	0.788

Chemical Cells with Transference. In chemical cells with transference the emf again results from a chemical reaction occurring within the cell, but this time assembly of the electrodes leads to a liquid junction between solutions of different electrolytes. Cells of this type are very common; in fact, of those described heretofore, such cells as

[1] From MacInnes, "Principles of Electrochemistry," Reinhold Publishing Corporation, New York, 1939, p. 187.

$$\text{Zn} \mid \text{Zn}^{++} \parallel \text{Cd}^{++} \mid \text{Cd}$$
$$\text{Hg} \mid \text{Hg}_2\text{Cl}_2(\text{s}), \text{KCl(1N)} \parallel \text{Cu}^{++} \mid \text{Cu}$$
$$\text{Tl} \mid \text{Tl}^+ \parallel \text{Sn}^{++} \mid \text{Sn}$$

as well as $\quad \text{Pt} \mid \text{Sn}^{++++}, \text{Sn}^{++} \parallel \text{Ce}^{+++}, \text{Ce}^{++++} \mid \text{Pt}$
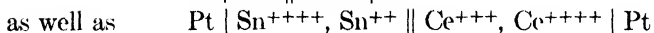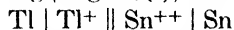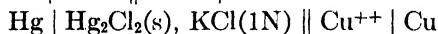
belong to this category. In treating these we have assumed that the total measured emf is the sum of the two electrode potentials. Although this *may* be the case to a fairly close approximation, it is not exactly true, and consequently chemical cells where transference is present require further consideration.

It has already been mentioned in connection with the Nernst solution tension theory that contact between two solutions of different concentrations, or different ions, or both, leads to a contact or junction potential $\mathcal{E}_j$. This potential arises, not from the reaction taking place in the cell, but from *diffusion* of ions across the boundary between the two solutions. Because of a concentration gradient existing across the boundary, ions tend to diffuse from the side of higher to that of lower concentration. Were the two ions of an electrolyte to migrate with equal velocities, such diffusion would cause no complications. However, as in general this does not occur, the faster ion moves across the boundary ahead of the slower, and a *separation of charges*, analogous to a Helmholtz double layer, results. Since every double layer involves a potential difference across it, this separation of charges causes the establishment of a junction potential which is measured experimentally along with the two electrode potentials, and appears in the total emf of the cell. In other words, whenever a junction potential $\mathcal{E}_j$ is present, the emf of a cell is no longer $\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2$, where $\mathcal{E}_1$ and $\mathcal{E}_2$ are the two electrode potentials, but instead

$$\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_j \tag{75}$$

In most instances junction potentials cannot be measured separately. Consequently various attempts have been made by Henderson, Planck, and others [1] to calculate junction potentials, and to arrive thereby through equation (75) at the sum of the electrode potentials corresponding to the cell reaction. These attempts all lead to complex expressions difficult to apply. The only simple case arises when the two electrolytes in contact are both of the 1–1 valence type, at the same concentration, and possess an ion in common; i.e., when the junctions are, say,

$$\text{HCl(0.1 N)} \mid \text{KCl(0.1 N)}$$
$$\text{KCl(0.05 N)} \mid \text{NaCl(0.05 N)}$$
$$\text{NaCl(0.01 N)} \mid \text{NaNO}_3\text{(0.01 N)}$$
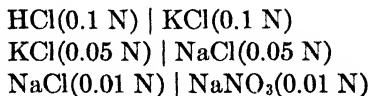
[1] See MacInnes, *op. cit.*, Chapter XIII.

For these the junction potential can be calculated quite satisfactorily by the Lewis and Sargent [1] relation

$$\mathcal{E}_j = \frac{RT}{\mathcal{F}} \ln \frac{\Lambda_1}{\Lambda_2} \tag{76}$$

where Λ_1 and Λ_2 are the equivalent conductances of the two solutions in contact. In most other instances, however, the progress made is not sufficient to permit the calculation of $\mathcal{E}_1 + \mathcal{E}_2$ from $\mathcal{E}$ without some ambiguity. For this reason chemical cells with transference are not considered suitable for *exact* evaluation of the thermodynamic properties of such cells. Nevertheless, under suitable experimental conditions these cells can be utilized to yield valuable even though not completely exact information, and they are employed a great deal for measurements of pH, electrometric titrations, etc.

Methods of Forming Liquid Junctions. The simplest means of forming a liquid junction between two solutions is to bring the two together in a tube without causing undue mixing. Although when carefully prepared such *static junctions* may remain sharp and give reproducible potentials, a more satisfactory arrangement is the *flowing junction.* In the latter the two solutions are brought together as two streams
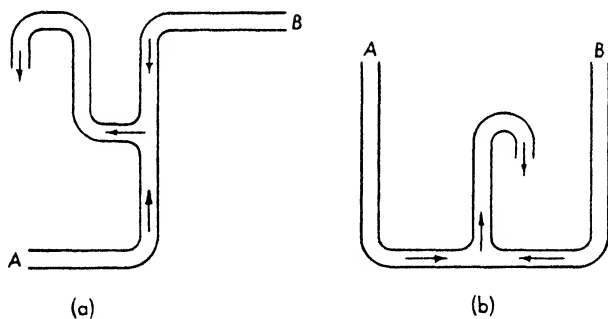
Fig. 8. Flowing Junctions

which merge in a common tube and flow out. Two simple forms of flowing junctions are shown in Fig. 8. In both Figs. (8a) and (8b) A is the entry tube for one solution, B for the other, while the third tube is the one in which the two solutions contact each other and pass out of the apparatus. By this uncomplicated artifice the boundary between the two solutions can be kept continually renewed and sharp.

However, the type of junction employed most frequently is the *salt bridge.* In this scheme a fairly concentrated solution of a salt, usually 1 N or saturated potassium chloride, is interposed between the two solutions in the manner indicated in Fig. 5. This salt bridge is supposed to

---

[1] Lewis and Sargent, J. Am. Chem. Soc., **31**, 363 (1909).

minimize or *reduce* the junction potential. Exactly how this reduction of the junction potential is accomplished is not quite clear, but it is supposed to be associated with the fact that the two ions in potassium chloride possess about equal velocities, and these operate to yield junction potentials between the two solutions and the bridge which are opposite in sign, and hence cancel each other to a degree. Whether such a reduction in junction potentials with the aid of salt bridges actually occurs is problematical. Still more questionable is it whether such bridges can reduce the junction potentials to a point where they are negligible. Nevertheless, effective or not, the fact remains that salt bridges have come into extensive use, and are encountered in various electrochemical assemblies.

Calculations Involving Chemical Cells with Transference. To illustrate calculations involving these cells we may assume that a salt bridge eliminates completely the junction potential, and hence that the measured cell emf is still the sum of the two electrode potentials. When $\mathcal{E}_j$ is disregarded in this manner, it is customary to write two vertical lines between the two solutions, namely,

$$\text{Zn} \mid \text{Zn}^{++} \parallel \text{Cd}^{++} \mid \text{Cd}$$
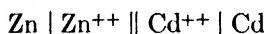
On the other hand, when the presence of the liquid junction is recognized and taken into account, only one line is interposed, as

$$\text{Zn} \mid \text{Zn}^{++} \mid \text{Cd}^{++} \mid \text{Cd}$$
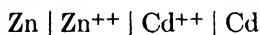
It is characteristic of chemical cells with transference that they involve in their emf equations the *activities of ions* rather than mean activities of the electrolyte as a whole. This would suggest that such cells may be suitable for the evaluation of activities, and hence the activity coefficients, of individual ions. However, this is not quite the case because of the junction potential uncertainty. In fact, it is quite generally agreed that there is no exact thermodynamic method known at present by which anything more than the geometric mean activity coefficient of an electrolyte can be determined. Therefore, when the emf's of chemical cells with transference are used to estimate activity coefficients of ions, this is done with the understanding that the quantities evaluated possess only quasi-thermodynamic significance. Again, in calculating the emf's of such cells from molalities and activity coefficients, the usual practice is to assume that the activity coefficient of any ion is essentially equal to the geometric mean activity coefficient of the electrolyte, and to use the latter.

With these remarks in mind we are ready to consider a specific calculation. Suppose that it is desired to estimate at 25° C the emf of the cell

$$\text{Zn} \mid \text{ZnCl}_2(m = 0.5) \parallel \text{CdSO}_4(m = 0.1) \mid \text{Cd}$$
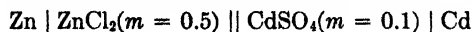
where the m's are the molalities of the two solutions. For this cell the emf is given by

$$\mathcal{E} = \mathcal{E}^0 - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_{Zn^{++}}}{a_{Cd^{++}}}$$

Since $a_{Zn^{++}} = m_{Zn^{++}}\gamma_{Zn^{++}}$, and $a_{Cd^{++}} = m_{Cd^{++}}\gamma_{Cd^{++}}$, where the γ's are the activity coefficients of the ions in their respective solutions, the above equation may be written as

$$\mathcal{E} = \mathcal{E}^0 - \frac{RT}{2\,\mathcal{F}} \ln \frac{m_{Zn^{++}}\gamma_{Zn^{++}}}{m_{Cd^{++}}\gamma_{Cd^{++}}}$$

Now, at 25° C $\mathcal{E}^0 = 0.3590$ volt for this cell. Again, for 0.5 m $ZnCl_2$ $\gamma = 0.376$, while for 0.1 m $CdSO_4$ $\gamma = 0.137$. Assuming that these γ's are also the ionic activity coefficients, we find from these quantities and $m_{Zn^{++}} = 0.5$, $m_{Cd^{++}} = 0.1$, for $\mathcal{E}$:

$$\mathcal{E} = 0.3590 - \frac{0.05915}{2} \log_{10} \frac{(0.5)(0.376)}{(0.1)(0.137)}$$
$$= 0.3590 - 0.0336$$
$$= 0.3254 \text{ volt at } 25° \text{ C}$$

Concentration Cells. Unlike chemical cells, in which the emf arises from a chemical reaction, *concentration cells depend for their emf on a transfer of material from one electrode to the other due to a concentration difference between the two.* This difference in concentration may arise from the fact that two identical electrodes dipping in the same solution may be at two different concentrations, as two hydrogen electrodes at two unlike gas pressures immersed in the same solution of hydrogen ions, namely,

$$H_2(P_{H_2} = P_1) \mid H^+ \mid H_2(P_{H_2} = P_2)$$
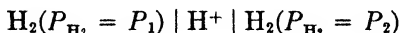
or two amalgam electrodes at two different concentrations may dip in a solution of the metal ions, as

$$Cd(Hg)(C_{Cd} = C_1) \mid Cd^{++} \mid Cd(Hg)(C_{Cd} = C_2)$$
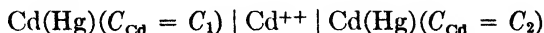
Again, the difference in concentration may not be in the electrodes, but in the solutions with which they are in contact, as for instance in the cells

$$H_2(g, 1 \text{ atm}) \mid H^+(a_1) \mid H^+(a_2) \mid H_2(g, 1 \text{ atm})$$
$$Ag \mid Ag^+(a_1) \mid Ag^+(a_2) \mid Ag$$
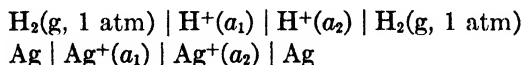
In the first two cells mentioned there is no liquid junction present, and hence they are *concentration cells without transference.* On the other hand, the latter two cells involve a liquid junction between two solutions of the same kind but unlike concentrations, and therefore these constitute *concentration cells with transference.*

Electrode Concentration Cells without Transference. To see why an emf arises in cells where two like electrodes are at different con-

centrations, but the electrolyte is the same for both, let us take first the cell

$$H_2(P_{H_2} = P_1) \mid H^+(a_{H^+}) \mid H_2(P_{H_2} = P_2) \qquad (77)$$
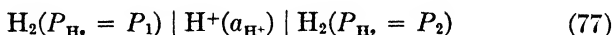
For the electrode on the left, the oxidation reaction yields

$$\frac{1}{2} H_2(P_1) = H^+(a_{H^+}) + \ominus \qquad (78a)$$
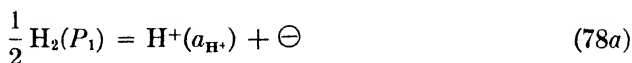
and therefore, since $\mathcal{E}^0_{H_2} = 0$,

$$\mathcal{E}_1 = -\frac{RT}{\mathcal{F}} \ln \frac{a_{H^+}}{P_1^{1/2}} \qquad (78b)$$

Similarly, for the reduction at the right-hand electrode, the process is

$$H^+(a_{H^+}) + \ominus = \frac{1}{2} H_2(P_2) \qquad (79a)$$
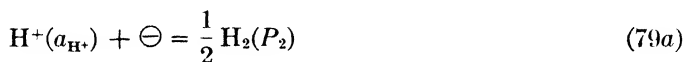
and

$$\mathcal{E}_2 = -\frac{RT}{\mathcal{F}} \ln \frac{P_2^{1/2}}{a_{H^+}} \qquad (79b)$$

On adding equations (78a) and (79a), the cell reaction follows as

$$\frac{1}{2} H_2(P_1) = \frac{1}{2} H_2(P_2) \qquad (80a)$$

while the cell emf, on adding equations (78b) and (79b), is

$$\begin{aligned}
\mathcal{E} &= -\frac{RT}{\mathcal{F}} \ln \frac{a_{H^+}}{P_1^{1/2}} - \frac{RT}{\mathcal{F}} \ln \frac{P_2^{1/2}}{a_{H^+}} \\
&= -\frac{RT}{\mathcal{F}} \ln \frac{P_2^{1/2}}{P_1^{1/2}} \\
&= -\frac{RT}{2\mathcal{F}} \ln \frac{P_2}{P_1} \qquad (80b)
\end{aligned}$$

Equation (80a) shows that the cell reaction involves merely the *transfer of 0.5 mole of hydrogen gas from a pressure P_1 atmospheres at one electrode to a pressure P_2 at the other;* i.e., that the cell reaction for the spontaneous process is an expansion of hydrogen gas from a pressure P_1 to a pressure P_2. Again, equation (80b) shows that the emf resulting from this expansion depends only on the two pressures, and is completely *independent of the activity of the hydrogen ions* in which the electrodes are immersed.

Consider next another cell of this type, namely,

$$Zn(Hg)(a_{Zn} = a_1) \mid Zn^{++}(a_{Zn^{++}}) \mid Zn(Hg)(a_{Zn} = a_2) \qquad (81)$$
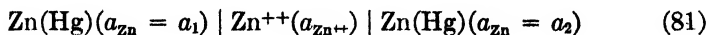
consisting of two zinc amalgams of activities of zinc equal to a_1 and a_2 immersed in a solution of zinc ions of activity $a_{Zn^{++}}$. Analogous to equations (39) and (40), the electrode reaction on the left is

$$Zn(Hg)(a_1) = Zn^{++}(a_{Zn^{++}}) + 2 \ominus \qquad (82a)$$
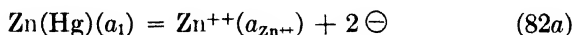
with $\mathcal{E}_1$ given by

$$\mathcal{E}_1 = \mathcal{E}_{Zn}^0 - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_{Zn^{++}}}{a_1} \qquad (82b)$$

Again, for the reduction on the right, we have

$$Zn^{++}(a_{Zn^{++}}) + 2 \ominus = Zn(Hg)(a_2) \qquad (83a)$$
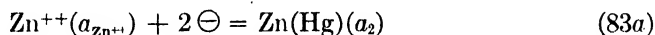
and

$$\mathcal{E}_2 = \mathcal{E}_{Zn}^{0\prime} - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_2}{a_{Zn^{++}}} \qquad (83b)$$

where $\mathcal{E}_{Zn}^{0\prime} = - \mathcal{E}_{Zn}^0$. We obtain, therefore, for the cell reaction,

$$Zn(Hg)(a_1) = Zn(Hg)(a_2) \qquad (84a)$$
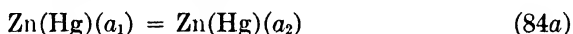
and for the cell emf,

$$
\begin{aligned}
\mathcal{E} &= \mathcal{E}_1 + \mathcal{E}_2 \\
&= \mathcal{E}_{Zn}^0 - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_{Zn^{++}}}{a_1} + (- \mathcal{E}_{Zn}^0) - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_2}{a_{Zn^{++}}} \\
&= - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_2}{a_1} \qquad (84b)
\end{aligned}
$$

From equation (84a) it is evident that the emf of this amalgam concentration cell arises from a transfer of zinc from the amalgam where its activity is a_1 to the amalgam where its activity is a_2. Further, equation (84b) shows that this emf depends only on the ratio of the activities of the zinc in the two amalgams, and not at all on the activity of the zinc ions in the solution. In the final cell emf equation $\mathcal{E}^0$ does not appear. This is true of all concentration cells. We may conclude, therefore, *that for concentration cells $\mathcal{E}^0$ is zero*, and the Nernst equation takes on the simplified form

$$\mathcal{E}_{cell} = - \frac{RT}{n\,\mathcal{F}} \ln \frac{a_2}{a_1} \qquad (85)$$

Since in cells of the type under discussion a_2 and a_1 refer to the activities of the metal in the amalgams, we may assume that for dilute amalgams these do not differ essentially from the concentrations C_1 and C_2. Hence equation (84b) may be written approximately as

$$\mathcal{E} = - \frac{RT}{2\,\mathcal{F}} \ln \frac{C_2}{C_1} \qquad (86)$$

The applicability of the latter equation to zinc amalgams has been confirmed by Meyer.[1]

Electrolyte Concentration Cells without Transference. Consider a chemical cell without transference, such as

$$H_2(g, 1 \text{ atm}) \mid HCl(a_1), AgCl(s) \mid Ag \tag{87a}$$
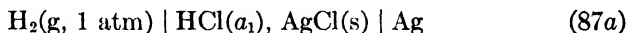
For this cell we have seen that the cell reaction is

$$\frac{1}{2} H_2(g, 1 \text{ atm}) + AgCl(s) = Ag(s) + HCl(a_1) \tag{87b}$$
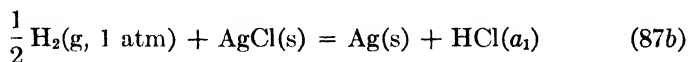
and the cell emf, equation (71),

$$\mathcal{E}_1 = \mathcal{E}^0 - \frac{RT}{\mathcal{F}} \ln a_1 \tag{87c}$$

Again, for the same cell but with a different activity of hydrochloric acid, namely,

$$H_2(g, 1 \text{ atm}) \mid HCl(a_2), AgCl(s) \mid Ag \tag{88a}$$

the cell reaction will be

$$\frac{1}{2} H_2(g, 1 \text{ atm}) + AgCl(s) = Ag(s) + HCl(a_2) \tag{88b}$$
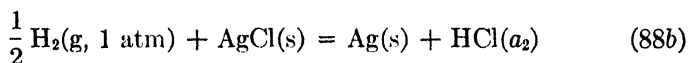
and the cell emf will be

$$\mathcal{E}_2 = \mathcal{E}^0 - \frac{RT}{\mathcal{F}} \ln a_2 \tag{88c}$$

If these two cells are connected together so as to *oppose* each other, i.e.,

$$\frac{1}{2} H_2(g,1\text{atm}) \mid HCl(a_1), AgCl(s) \mid Ag\text{-}Ag \mid AgCl(s), HCl(a_2) \mid H_2(g,1\text{atm}) \tag{89a}$$
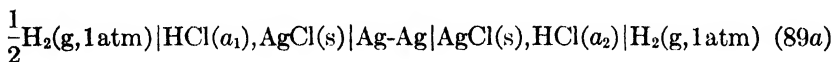
the overall reaction of the combination will be the *difference* between equations (87b) and (88b), namely,

$$\frac{1}{2} H_2(g, 1 \text{ atm}) + AgCl(s) - \frac{1}{2} H_2(g, 1 \text{ atm}) - AgCl(s) =$$
$$Ag(s) + HCl(a_1) - Ag(s) - HCl(a_2)$$
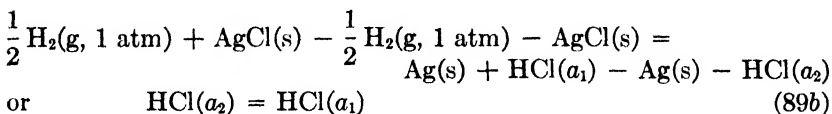
or $\qquad HCl(a_2) = HCl(a_1) \tag{89b}$

Similarly the emf of (89a) will be the *difference* between equations (87c) and (88c). Hence,

$$\mathcal{E} = \mathcal{E}_1 - \mathcal{E}_2$$
$$= \left(\mathcal{E}^0 - \frac{RT}{\mathcal{F}} \ln a_1\right) - \left(\mathcal{E}^0 - \frac{RT}{\mathcal{F}} \ln a_2\right)$$
$$= -\frac{RT}{\mathcal{F}} \ln \frac{a_1}{a_2} \tag{89c}$$

---

[1] Meyer, G., Zeit. physik. Chem., **7**, 477 (1891).

According to equation (89b) the overall reaction resulting from the combination given in (89a) is a *transfer* for each faraday drawn from the cell of 1 mole of hydrochloric acid from the solution where the activity is a_2 to the solution where the activity is a_1. Consequently, whereas each of the individual cells constituting (89a) is a chemical cell, the combination of the two is a *concentration cell without transference in which the emf arises from different concentrations of the electrolyte.* For this emf to be positive, equation (89c) shows that a_2 must be greater than a_1. In other words, the transfer process is spontaneous for passage of electrolyte from the more concentrated to the more dilute solution, as is to be expected.

Concentration cells of this type are not limited to the one described, but may be assembled from any chemical cells without transference. For instance, a sodium chloride concentration cell results from the combination

$$Na(Hg) \mid NaCl(a_1), AgCl(s) \mid Ag-Ag \mid AgCl(s), NaCl(a_2) \mid Na(Hg)$$
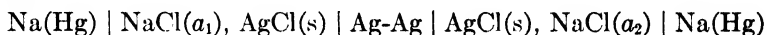
a zinc sulfate concentration cell from

$$Zn(Hg) \mid ZnSO_4(a_1), PbSO_4(s) \mid Pb(Hg)-Pb(Hg) \mid PbSO_4(s),$$
$$ZnSO_4(a_2) \mid Zn(Hg)$$
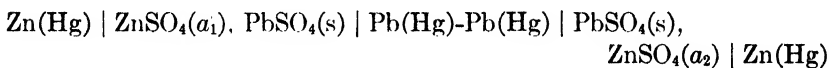
and a potassium bromide concentration cell from

$$K(Hg) \mid KBr(a_1), AgBr(s) \mid Ag-Ag \mid AgBr(s), KBr(a_2) \mid K(Hg)$$
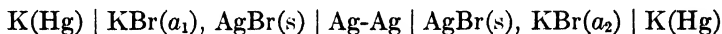
Such cells are suitable for the determination of the activity coefficients of the electrolytes involved. For this purpose the usual practice is to keep one of the electrolyte concentrations constant, and to measure the emf of the cells with various other concentrations of the second solution. If for a_1 and a_2 in equation (89c) we make the substitutions $a_1 = m_1^2\gamma_1^2$ and $a_2 = m_2^2\gamma_2^2$, the expression for $\mathcal{E}$ becomes

$$
\begin{aligned}
\mathcal{E} &= -\frac{RT}{\mathcal{F}} \ln \frac{m_1^2\gamma_1^2}{m_2^2\gamma_2^2} \\
&= \frac{2\,RT}{\mathcal{F}} \ln \frac{m_2\gamma_2}{m_1\gamma_1} \\
&= \frac{2\,RT}{\mathcal{F}} \ln \frac{m_2}{m_1} + \frac{2\,RT}{\mathcal{F}} \ln \frac{\gamma_2}{\gamma_1}
\end{aligned}
\tag{90}
$$

As the molalities are known, equation (90) permits calculating the ratio of the activity coefficients of the electrolyte at any molality m_2 to that at the reference concentration m_1 from the measured emf's. Hence when γ_1 corresponding to molality m_1 is available, this yields immediately the various activity coefficients.

Concentration Cells with Transference. A typical concentration cell with transference is

$$H_2(g, 1 \text{ atm}) \mid HCl(a_1) \mid HCl(a_2) \mid H_2(g, 1 \text{ atm}) \qquad (91)$$
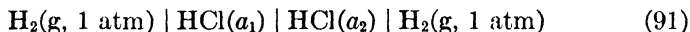
consisting of two identical hydrogen electrodes immersed in two hydrochloric acid solutions of different concentrations. For this cell the total emf is evidently composed of the two single electrode potentials and the potential at the junction, $\mathcal{E}_j$. Further, the overall cell reaction is the sum of the two electrode processes and any material transfer taking place across the junction. In a case such as this, where the electrolytes constituting the junction are identical except for their concentrations, these various processes can be analyzed, and an equation for the cell can be arrived at which takes into account the junction potential.

To understand how this is possible, assume that the electrode on the left in equation (91) is negative, which will be true when $a_2 > a_1$. For a faraday of electricity drawn from this cell, the reaction at the negative electrode will be

$$\frac{1}{2} H_2(g, 1 \text{ atm}) = H^+(a_1) + \ominus$$

that at the positive,

$$H^+(a_2) + \ominus = \frac{1}{2} H_2(g, 1 \text{ atm})$$
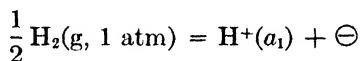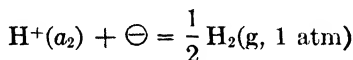
and hence the sum of the two electrode reactions is

$$H^+(a_2) = H^+(a_1) \qquad (92a)$$

However, when electrons flow externally from left to right, they must complete the circuit by passing *through the cell from right to left;* i.e., electrons must pass across the liquid junction from right to left. This electron current in the cell is composed, of course, not of free electrons, but of negative ions, namely Cl^-, moving from *right to left*, and positive ions, or H^+, moving across the junction from *left to right*. If we let now t_- be the transference number of the chloride ions, then for every faraday passing through the cell t_- equivalents of chloride ions will be transported from the solution where the activity is a_2, to the solution where the activity is a_1. Or,

$$t_-Cl^-(a_2) = t_-Cl^-(a_1) \qquad (92b)$$

Again, $t_+ = 1 - t_-$ equivalents of hydrogen ions will be transferred from the solution of activity a_1 to the solution of activity a_2, namely,

$$(1 - t_-)H^+(a_1) = (1 - t_-)H^+(a_2) \qquad (92c)$$
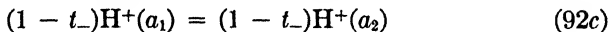
Therefore, in order to obtain the *net transfer* of material we must add equations (92b) and (92c) to (92a). We receive thus, for the overall process of the cell,

$$H^+(a_2)+t_-Cl^-(a_2)+(1-t_-)H^+(a_1) = H^+(a_1)+t_-Cl^-(a_1)+(1-t_-)H^+(a_2)$$
$$t_-H^+(a_2) + t_-Cl^-(a_2) = t_-H^+(a_1) + t_-Cl^-(a_1)$$
$$t_-HCl(a_2) = t_-HCl(a_1) \tag{93}$$
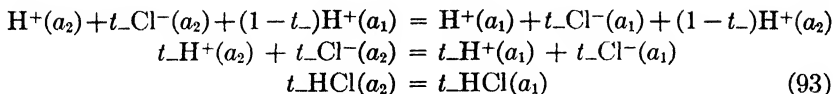
Equation (93) shows that in the concentration cell with transference t_- equivalents of hydrochloric acid are transferred from the solution of activity a_2 to the solution of activity a_1 for every faraday of electricity. This may be contrasted with the same process in a concentration cell without transference, equation (89b), where for 1 faraday the transfer of a full equivalent of the electrolyte is accomplished.

Application of the Nernst emf equation to the process in (93) gives immediately for the cell emf

$$\begin{aligned}
\mathcal{E} &= -\frac{RT}{\mathcal{F}} \ln \frac{a_1^{t_-}}{a_2^{t_-}} \\
&= -\frac{t_- RT}{\mathcal{F}} \ln \frac{a_1}{a_2} \\
&= \frac{t_- RT}{\mathcal{F}} \ln \frac{a_2}{a_1}
\end{aligned} \tag{94a}$$

On insertion into this expression of the molalities and the activity coefficients, this emf becomes

$$\begin{aligned}
\mathcal{E} &= \frac{t_- RT}{\mathcal{F}} \ln \frac{m_2^2 \gamma_2^2}{m_1^2 \gamma_1^2} \\
&= \frac{2 t_- RT}{\mathcal{F}} \ln \frac{m_2 \gamma_2}{m_1 \gamma_1}
\end{aligned} \tag{94b}$$

and hence it is evident that the emf of such a cell can be calculated from a knowledge of the molalities, activity coefficients, and the *transport number of the ion other than the one to which the electrodes are reversible.*

Equations analogous to equations (93), (94a), and (94b) are applicable to any concentration cell with transference in which the electrodes are reversible to the *cation.* However, when the electrodes are reversible to the *anion,* as in the cell

$$Ag \mid AgCl(s), HCl(a_1) \mid HCl(a_2), AgCl(s) \mid Ag \tag{95}$$
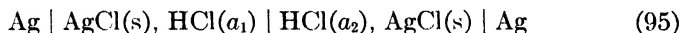
it can readily be deduced by repetition of the above reasoning that the cell reaction is

$$t_+HCl(a_1) = t_+HCl(a_2) \tag{96}$$

and the emf is given by

$$\begin{aligned}
\mathcal{E} &= -\frac{RT}{\mathcal{F}} \ln \frac{a_2^{t_+}}{a_1^{t_+}} \\
&= \frac{t_+ RT}{\mathcal{F}} \ln \frac{m_1^2 \gamma_1^2}{m_2^2 \gamma_2^2} \\
&= \frac{2 t_+ RT}{\mathcal{F}} \ln \frac{m_1 \gamma_1}{m_2 \gamma_2}
\end{aligned} \tag{97}$$

In this case evidently a_1 must be greater than a_2 for the reaction to be spontaneous.

The Junction Potential. The emf given by equation (94*b*) for cell (91) is the sum of the two electrode potentials and $\mathcal{E}_j$. It is of interest to ascertain how much of this total emf is due to the electrodes and how much to the liquid junction.

According to equation (92*a*), the sum of the two electrode reactions in cell (91) is

$$H^+(a_2) = H^+(a_1)$$

and hence the sum of the two electrode potentials, $\mathcal{E}_1 + \mathcal{E}_2$, is

$$
\begin{aligned}
\mathcal{E}_1 + \mathcal{E}_2 &= -\frac{RT}{\mathcal{F}} \ln \frac{(a_{H^+})_1}{(a_{H^+})_2} \\
&= \frac{RT}{\mathcal{F}} \ln \frac{(m_{H^+}\gamma_{H^+})_2}{(m_{H^+}\gamma_{H^+})_1}
\end{aligned}
\tag{98}
$$

If equation (98) is subtracted from (94*b*), the result is $\mathcal{E}_j$, namely,

$$
\begin{aligned}
\mathcal{E}_j &= \mathcal{E} - (\mathcal{E}_1 + \mathcal{E}_2) \\
&= \frac{2\,t_-RT}{\mathcal{F}} \ln \frac{m_2\gamma_2}{m_1\gamma_1} - \frac{RT}{\mathcal{F}} \ln \frac{(m_{H^+}\gamma_{H^+})_2}{(m_{H^+}\gamma_{H^+})_1}
\end{aligned}
$$

But $(m_{H^+})_2 = m_2$ and $(m_{H^+})_1 = m_1$. Further, if, as before, we take the activity coefficient of an ion equal to the mean activity coefficient of the electrolyte, $(\gamma_{H^+})_2$ becomes γ_2 and $(\gamma_{H^+})_1$ becomes γ_1. $\mathcal{E}_j$ follows thus as,

$$
\begin{aligned}
\mathcal{E}_j &= \frac{2\,t_-RT}{\mathcal{F}} \ln \frac{m_2\gamma_2}{m_1\gamma_1} - \frac{RT}{\mathcal{F}} \ln \frac{m_2\gamma_2}{m_1\gamma_1} \\
&= (2\,t_- - 1)\frac{RT}{\mathcal{F}} \ln \frac{m_2\gamma_2}{m_1\gamma_1}
\end{aligned}
\tag{99a}
$$

Alternately, since $(t_+ + t_-) = 1$, $(2\,t_- - 1) = (2\,t_- - t_+ - t_-) = (t_- - t_+)$, and equation (99*a*) may also be written as

$$
\mathcal{E}_j = (t_- - t_+)\frac{RT}{\mathcal{F}} \ln \frac{m_2\gamma_2}{m_1\gamma_1}
\tag{99b}
$$

We see, therefore, that besides the activities of the two solutions constituting the junction, $\mathcal{E}_j$ is determined also by the difference between the two transport numbers of the electrolyte. When these are equal, $(t_- - t_+) = 0$, and so is the junction potential. This fact is the basis for the use of potassium chloride as a bridge in cell measurements involving liquid junction, for in this electrolyte the two transference numbers are very nearly identical. In turn, when $t_- > t_+$, $\mathcal{E}_j$ is positive and adds to

the $\mathcal{E}_1 + \mathcal{E}_2$, while when $t_- < t_+$, $\mathcal{E}_j$ is negative and operates to yield for $\mathcal{E}$ of the cell a potential less than $\mathcal{E}_1 + \mathcal{E}_2$.

Similar considerations, applied to concentration cells with transference containing electrodes reversible to the *anion*, give for $\mathcal{E}_j$, instead of equation (99a)

$$\mathcal{E}_j = (2\,t_+ - 1)\,\frac{RT}{\mathcal{F}}\,\ln\frac{m_1\gamma_1}{m_2\gamma_2} \tag{100a}$$

or

$$\mathcal{E}_j = (t_+ - t_-)\,\frac{RT}{\mathcal{F}}\,\ln\frac{m_1\gamma_1}{m_2\gamma_2} \tag{100b}$$

Here m_1 is the more concentrated solution. From equation (100b) it is clear that again for $t_+ = t_-$ $\mathcal{E}_j = 0$, but this time $\mathcal{E}_j$ is positive for $t_+ > t_-$, and negative for $t_+ < t_-$.

Transference Numbers from Emf Measurements. Since concentration cells with transference involve transport numbers in their cell and emf equations, such cells may be utilized for estimating these quantities from emf data. For this purpose equations (94b) and (97) show that we need the emf's of the respective cells and the activity coefficients of the electrolytes used. However, it is possible to arrive at t_+ and t_- from emf's alone without a knowledge of the activity coefficients.

For any concentration cell *with* transference in which the electrodes are reversible to the *cation*, as

$$\text{M} \mid \text{MX}(a_1) \mid \text{MX}(a_2) \mid \text{M} \tag{101}$$

equation (94a) gives for the emf

$$\mathcal{E}_{\text{M}} = \frac{t_-RT}{\mathcal{F}}\,\ln\frac{a_2}{a_1} \tag{94a}$$

Again, for a concentration cell involving the same concentrations of the electrolyte but *without* transference, namely,

$$\text{M} \mid \text{MX}(a_1) \mid \text{X} - \text{X} \mid \text{MX}(a_2) \mid \text{M} \tag{102}$$

the reaction would be, analogous to equation (89b),

$$\text{MX}(a_2) = \text{MX}(a_1)$$
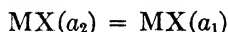
and the emf $\mathcal{E}$, according to equation (89c), would be

$$\mathcal{E} = \frac{RT}{\mathcal{F}}\,\ln\frac{a_2}{a_1} \tag{103}$$

On dividing equation (94a) by equation (103), we find that

$$\frac{\mathcal{E}_{\text{M}}}{\mathcal{E}} = \frac{\dfrac{t_-RT}{\mathcal{F}}\,\ln\dfrac{a_2}{a_1}}{\dfrac{RT}{\mathcal{F}}\,\ln\dfrac{a_2}{a_1}} = t_- \tag{104}$$

and hence t_- follows as the *ratio of the emf's of the cell with transference and the same cell without transference*. Similarly, for a concentration cell with transference reversible to the *anion*, as

$$X \mid MX(a_1) \mid MX(a_2) \mid X \tag{105}$$
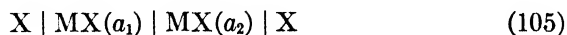
equation (97) yields for the emf,

$$\mathcal{E}_X = \frac{t_+ RT}{\mathcal{F}} \ln \frac{a_1}{a_2} \tag{97}$$

while with the same cell without transference, i.e.,

$$X \mid MX(a_1) \mid M - M \mid MX(a_2) \mid X \tag{106}$$

the emf is

$$\mathcal{E} = \frac{RT}{\mathcal{F}} \ln \frac{a_1}{a_2} \tag{107}$$

where $a_1 > a_2$. From the ratio of equation (97) to equation (107), t_+ follows then as

$$\frac{\mathcal{E}_X}{\mathcal{E}} = \frac{\dfrac{t_+ RT}{\mathcal{F}} \ln \dfrac{a_1}{a_2}}{\dfrac{RT}{\mathcal{F}} \ln \dfrac{a_1}{a_2}} = t_+ \tag{108}$$

Consequently, by measuring $\mathcal{E}_M$ of cell (101), $\mathcal{E}_X$ of cell (105), and $\mathcal{E}$ of cells (102) *or* (106), since these have the same emf, the transport numbers of the electrolyte MX can be obtained directly and independently from equations (104) and (108).

Solubility Products and Emf. The saturation solubility of any difficultly soluble salt such as silver bromide is represented by the equation

$$AgBr(s) = Ag^+ + Br^- \tag{109}$$
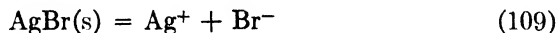
where the product of the activities of the two ions is a constant at any given temperature equal to the solubility product. Since this solubility product K_s is an equilibrium constant, it must be related by equation (36) to the $\mathcal{E}^0$ value of any cell in which the above reaction occurs. Hence, by finding the appropriate cell and its $\mathcal{E}^0$, it should be possible to calculate from the latter the solubility product.

Inspection of equation (109) suggests that the cell in which this reaction takes place is

$$Ag \mid AgBr(a_x), \ AgBr(s) \mid Ag \tag{110a}$$
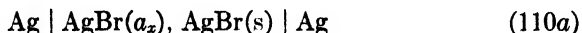
i.e., a cell composed of silver and silver-silver bromide electrodes immersed in a solution saturated with the salt. For the silver electrode, the reaction is

$$Ag(s) = Ag^+ + \ominus$$

for the silver-silver bromide electrode it is

$$AgBr(s) + \ominus = Ag(s) + Br^-$$
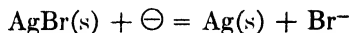
and therefore the sum of the two, namely,

$$AgBr(s) = Ag^+ + Br^- \tag{110b}$$
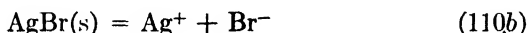
is identical with equation (109). From Table 3, $\mathcal{E}^0$ for the silver electrode at 25° C is seen to be $\mathcal{E}^0_{Ag} = -0.7990$ volt, $\mathcal{E}^0$ of the silver-silver bromide electrode for reduction $\mathcal{E}^0_{Ag-AgBr} = +0.0711$ volt, and hence $\mathcal{E}^0$ for cell (110a) follows as

$$\begin{aligned}
\mathcal{E}^0 &= \mathcal{E}^0_{Ag} + \mathcal{E}^0_{Ag-AgBr} \\
&= -0.7990 + 0.0711 \\
&= -0.7279 \text{ volt at } 25° \text{ C}
\end{aligned}$$

Substituting this value of $\mathcal{E}^0$ into equation (36), we obtain for the solubility product of silver bromide

$$\mathcal{E}^0 = \frac{RT}{n\mathcal{F}} \ln K_s$$
$$-0.7279 = .05915 \log_{10} K_s$$
$$K_s = 4.9 \times 10^{-13}$$

This value of K_s may be compared with $K_s = 7.7 \times 10^{-13}$ given in Table 11 of Chapter XVI. In a like manner may be calculated the solubility products of other difficultly soluble salts from tables of standard electrode potentials.

An alternate, but less exact, method of estimating the solubility product of a difficultly soluble salt depends on the measurement of the single electrode potential of an electrode reversible to one of the ions of the salt, and immersed in a solution saturated with this salt. To explain this method let us consider specifically the determination of the solubility product of silver chloride. For this purpose a solution of a chloride is taken, say 0.1 m potassium chloride, and the solution saturated with silver chloride by addition of a few drops of silver nitrate. A silver wire is inserted then into this solution, and the silver electrode thus formed is combined with a reference calomel electrode through a suitable bridge to form the cell

Ag | KCl(0.1 m, sat'd with AgCl) | Calomel Electrode
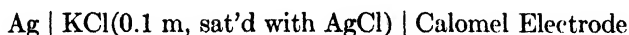
When the reference electrode is the 0.1 N calomel, the emf of this cell at 25° C is 0.0494 volt. Since $\mathcal{E}_C$ of the 0.1 N calomel for reduction is 0.3338 volt at 25° C, we have that

$$\mathcal{E} = \mathcal{E}_{Ag} + \mathcal{E}_C$$
$$0.0494 = \mathcal{E}_{Ag} + 0.3338$$
$$\mathcal{E}_{Ag} = -0.2844 \text{ volt}$$

i.e., the single electrode potential of the silver electrode in the 0.1 m potassium chloride saturated with silver chloride is -0.2844 volt at 25° C. But,

$$\mathcal{E}_{Ag} = \mathcal{E}_{Ag}^0 - \frac{RT}{\mathcal{F}} \ln a_{Ag^+}$$

and hence the activity of silver ions in this solution follows as

$$-0.2844 = -0.7990 - .0591 \log_{10} a_{Ag^+}$$

$$\log_{10} a_{Ag^+} = -\frac{0.5146}{.0591} = -8.708$$

$$a_{Ag^+} = 1.96 \times 10^{-9}$$

Now, for the activity of the chloride ions we may take the product of the molality, 0.1, and the mean activity coefficient of potassium chloride at this concentration, 0.770. We obtain, therefore, for K_s of silver chloride at 25° C,

$$K_s = a_{Ag^+} a_{Cl^-}$$
$$= (1.96 \times 10^{-9})(0.1 \times 0.770)$$
$$= 1.51 \times 10^{-10}$$

The thermodynamically exact value of K_s for this salt at 25° C is 1.76×10^{-10}.

Potentiometric Determination of pH. One of the most important and extensive applications of emf measurements is in determining the pH of various solutions. We have seen in Chapter XVI that indicators may be employed for this purpose, but the potentiometric or emf method constitutes a more fundamental and more accurate procedure. Further, colorimetric methods in the final analysis depend on the potentiometric method for the estimation of the pH of the buffers required, and hence at best these methods involve indirect rather than direct estimation of pH.

In all potentiometric pH work the procedure followed is very much the same. First a cell is assembled in which one of the electrodes is composed of an electrode reversible to hydrogen ions dipping into the solution whose pH is to be determined, while the other electrode is usually one of the calomels. Junction between the two is made either through

a salt bridge or by immersing the reference electrode directly into the solution. Next the emf of the combination, $\mathcal{E}$, is measured with a potentiometer. From $\mathcal{E}$ is subtracted the emf of the calomel to yield the single potential of the electrode reversible to hydrogen ions, and from the latter, in turn, the pH is calculated by means of the emf equation applicable to the particular electrode used.

Four electrodes are more or less suitable for potentiometric pH determinations. These are: (1) the hydrogen electrode, (2) the antimony electrode, (3) the quinhydrone electrode, and (4) the glass electrode. We turn now to a discussion of each of these.

The Hydrogen Electrode. This electrode is the standard of all pH measurements, and is the electrode against which all the others are checked. According to equation (49), the emf of this electrode is given by

$$\mathcal{E}_{H_2} = -\frac{RT}{\mathcal{F}} \ln a_{H^+} + \frac{RT}{\mathcal{F}} \ln P_{H_2}^{1/2} \tag{49}$$

Since by definition,

$$pH = -\log_{10} a_{H^+}$$

equation (49) in terms of pH is

$$\mathcal{E}_{H_2} = -\frac{2.3026\,RT}{\mathcal{F}} \log_{10} a_{H^+} + \frac{2.3026\,RT}{\mathcal{F}} \log_{10} P_{H_2}^{1/2}$$

$$= \frac{2.3026\,RT}{\mathcal{F}} pH + \frac{2.3026\,RT}{\mathcal{F}} \log_{10} P_{H_2}^{1/} \tag{111}$$

i.e., the emf of the hydrogen electrode is at any given temperature a function of the pH of the solution and the pressure of the hydrogen gas. When the latter is exactly 1 atmosphere, equation (111) reduces to

$$\mathcal{E}_{H_2} = \left(\frac{2.3026\,RT}{\mathcal{F}}\right) pH \tag{112}$$

and hence under these conditions *the emf of the hydrogen electrode is linearly dependent only on the pH of the solution.*

It was mentioned already that the hydrogen electrode consists of a piece of platinized platinum foil, immersed in the solution under test, around which is bubbled hydrogen gas. For satisfactory operation the platinized coat must be freshly deposited from a solution of chloroplatinic acid, and the hydrogen gas must be carefully purified to remove any impurities, particularly oxygen. With all details carefully controlled, this electrode can give very accurate results. It can be used over the entire pH range provided the solution does not contain oxidizing agents that may react with the hydrogen or metals that may be thrown out of

solution by the gas. However, because of the ease with which this electrode is poisoned even by oxygen of the air, and because of the rather elaborate setup required, its use is pretty much confined to the laboratory. This electrode lacks the portability, the ease of operation, the freedom from complicating factors, and the speed desired in routine industrial measurements.

To illustrate the calculation of pH from data obtained with a hydrogen electrode, let us take the following example. Suppose it is found that the emf of the cell

$$H_2 \mid \text{Solution (pH} = x) \mid 1 \text{ N calomel}$$
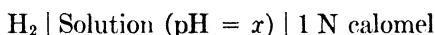
is $\mathcal{E} = 0.5164$ volt at 25° C when the corrected barometric pressure is 754.1 mm Hg. Subtracting from this emf the potential of the calomel electrode, namely, $\mathcal{E}_C = 0.2800$ volt at 25° C, we obtain for the emf of the hydrogen electrode

$$\begin{aligned} \mathcal{E}_{H_2} &= \mathcal{E} - \mathcal{E}_C \\ &= 0.5164 - 0.2800 \\ &= 0.2364 \text{ volt} \end{aligned}$$

This emf is due to both the pH of the solution and the pressure of the hydrogen gas. To find the value of the latter we must remember that the total barometric pressure at which the gas escapes from solution is composed of both the pressure of the hydrogen and the pressure of water vapor with which this gas becomes saturated in bubbling through the solution; i.e.,

$$P_{H_2} + P_{H_2O} = P_{\text{barometer}}$$

For P_{H_2O} we may take without any error the vapor pressure of water at 25° C, 23.8 mm. P_{H_2} is then

$$\begin{aligned} P_{H_2} &= P_{\text{barometer}} - P_{H_2O} \\ &= 754.1 - 23.8 \\ &= 730.3 \text{ mm Hg} = 0.961 \text{ atm} \end{aligned}$$

Substituting now $\mathcal{E}_{H_2} = 0.2364$ and $P_{H_2} = 0.961$ atmospheres into equation (111), the pH of the solution follows as

$$0.2364 = .0591 \text{ pH} + \frac{.0591}{2} \log_{10} 0.961$$

$$\text{pH} = \frac{0.2364 + 0.0005}{.0591}$$

$$= 4.01$$

It will be observed that the pressure correction to the emf is small, and amounts in this case to only 0.01 of a pH unit. Hence in any but the most precise work this correction can generally be disregarded.

The Antimony Electrode. According to the theory of the antimony-antimony trioxide electrode developed on p. 538 and epitomized in equation (61), this electrode should be reversible to hydrogen ions. It should be applicable, therefore, to pH determinations. Although this is approximately the case, and although this electrode is very convenient to use, it suffers from some limitations which greatly circumscribe its utility. In the first place, the electrode cannot be used below a pH of about 2, or above a pH of about 8. Outside these limits the electrode tends to dissolve to form on the acid side positive Sb^{+++} or SbO^+ ions, on the alkaline side antimonate ions. Again, it has been found that the emf of the electrode, even within the limits given, does not quite follow with pH the slope theoretically predicted by equation (61), i.e., 0.0591 at 25° C. Finally, at a pH of about 7 or 8 the electrode occasionally develops "blind spots," and tends to give erratic values for the emf. For these various reasons the antimony electrode is hardly ever used for pH measurements without previous calibration against a set of known buffers. However, as these shortcomings are not serious when the electrode is used in acid-base potentiometric titrations, the antimony electrode finds quite extensive application for this purpose.

The Quinhydrone Electrode. This interesting pH electrode consists of a platinum or gold wire immersed in the solution whose pH is to be determined, after the latter has been saturated with the difficultly soluble organic compound *quinhydrone*. Why such a system should be reversible to hydrogen ions may be gathered from the following argument.

When the organic compound *hydroquinone*, HO⟨◯⟩OH, is dissolved in water, it tends to undergo to a small extent oxidation to *quinone*, O=⟨◯⟩=O according to the reaction

$$\underset{\text{(HQ)}}{\overset{\text{OH}}{\underset{\text{OH}}{\bigcirc}}} = \underset{\text{(Q)}}{\overset{\text{O}}{\underset{\text{O}}{\bigcirc}}} + 2\,H^+ + 2\ominus \qquad (113)$$
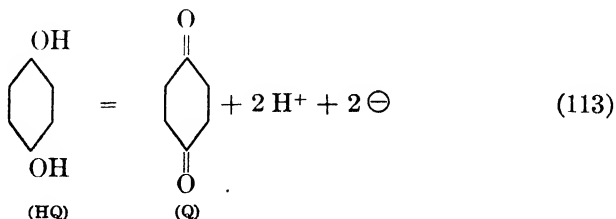
If an inert wire is inserted into this solution, it will pick up the electrons thus liberated and will acquire the potential $\mathcal{E}_Q$, corresponding to equation (113),

$$\mathcal{E}_Q = \mathcal{E}_Q^0 - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_{H^+}^2 a_Q}{a_{HQ}}$$

$$= \mathcal{E}_Q^0 - \frac{RT}{\mathcal{F}} \ln a_{H^+} - \frac{RT}{2\,\mathcal{F}} \ln \frac{a_Q}{a_{HQ}} \qquad (114)$$

where $\mathcal{E}_Q^0$ is the standard potential of this organic oxidation-reduction electrode, a_{H^+} the activity of the hydrogen ions in solution, and a_Q and a_{HQ} are the activities of the neutral quinone and hydroquinone. From equation (114) it is evident that $\mathcal{E}_Q$ would be a function of pH alone if only the last term could be disposed of by making $a_Q = a_{HQ}$. For this purpose we take advantage of the fact that quinone and hydroquinone form a double compound, *quinhydrone*, composed of one molecule each of quinone and hydroquinone. In aqueous solutions this compound is only difficultly soluble; but, whatever amount does dissolve dissociates to yield *equal quantities* of the two constituents. Since now quinone and hydroquinone are nonelectrolytes, and since their concentrations even in a saturated solution are at best small, we may suppose that not only are the concentrations of the two equal, but their activity coefficients as well. We may write thus $a_Q = a_{HQ}$, and hence for any solution *saturated with quinhydrone* equation (114) becomes

$$\mathcal{E}_Q = \mathcal{E}_Q^0 - \frac{RT}{\mathcal{J}} \ln a_{H^+}$$

$$= \mathcal{E}_Q^0 + \left(\frac{2.3026\ RT}{\mathcal{J}}\right) \text{pH} \tag{115}$$

Equation (115) gives the single electrode potential of the *quinhydrone electrode*. It indicates that $\mathcal{E}_Q$ under the conditions outlined *depends only on the pH of the solution*, and that the pH can be determined from the measured value of $\mathcal{E}_Q$ as soon as $\mathcal{E}_Q^0$ is available. Careful comparison of the quinhydrone and hydrogen electrodes has established that, between 0 and 40° C, $\mathcal{E}_Q^0$ follows the relation

$$\mathcal{E}_Q^0 = -\ 0.7176 + 0.000728\ t°\ \text{C} \tag{116}$$

which leads to $\mathcal{E}_Q^0 = -\ 0.6994$ volt at 25° C for the *oxidation* reaction given in equation (113). For this temperature we have, therefore,

$$\mathcal{E}_Q = -\ 0.6994 + .05915\ \text{pH} \tag{117}$$

The quinhydrone electrode when combined with a normal calomel electrode yields a cell in which up to a pH = 7.1 at 25° C the *quinhydrone electrode is positive*, i.e.,

$$\overset{-}{\text{1 N Calomel}} \left\|\begin{array}{l} \text{Solution (pH} = x), \\ \text{Sat'd quinhydrone} \end{array}\right| \text{Pt} \quad \text{pH} < 7.1 \tag{118a}$$

Consequently up to pH = 7.1 the calomel electrode undergoes an oxidation for which $\mathcal{E}_C = -\ 0.2800$ at 25° C, while the quinhydrone electrode undergoes a *reduction* with $\mathcal{E}_Q$ given by the reverse of equation (117). Adding $\mathcal{E}_C$ and $\mathcal{E}_Q$, we obtain for the emf, $\mathcal{E}$, of the combination,

$$\mathcal{E} = \mathcal{E}_Q + \mathcal{E}_C$$

$$= 0.6994 - 0.05915\ \text{pH} - 0.2800$$

$$= 0.4194 - 0.05915\ \text{pH} \tag{118b}$$

On the other hand, above pH = 7.1 the quinhydrone electrode becomes negative with respect to the normal calomel, i.e., the cell is now

$$\overset{-}{Pt} \left| \begin{array}{l} \text{Solution (pH = } x'\text{),} \\ \text{Sat'd quinhydrone} \end{array} \right\| 1 \text{ N Calomel} \qquad \text{pH} > 7.1 \quad (119a)$$
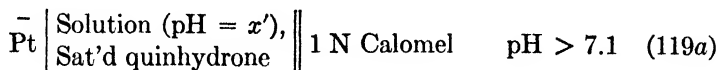
and the emf, $\mathcal{E}$, becomes the sum of equation (117) and $\mathcal{E}_C = 0.2800$, namely,

$$\begin{aligned} \mathcal{E} &= \mathcal{E}_Q + \mathcal{E}_C \\ &= -0.6994 + 0.05915 \text{ pH} + 0.2800 \\ &= -0.4194 + 0.05915 \text{ pH} \end{aligned} \quad (119b)$$

At pH = 7.1, $\mathcal{E} = 0$, and both equations apply. To facilitate the use of these equations we may readily combine them into the expression

$$\text{pH} = \frac{0.4194 - \mathcal{E}}{.05915} \quad (120)$$

where the experimentally observed value of the emf $\mathcal{E}$ of the cell is taken as *positive when the quinhydrone electrode is positive*, and *negative when the quinhydrone electrode is negative*.

Equations similar to (120) can readily be derived for the saturated or 0.1 N calomels used in connection with this electrode. In fact, all that need be done is to subtract from 0.6994 the value of the appropriate calomel electrode to yield the constant in the numerator of equation (120).

The quinhydrone electrode is very simple to use, and reaches equilibrium quite rapidly. It is highly satisfactory for pH determinations in solutions below pH = 8.5 provided no strong oxidizing or reducing agents are present to attack the quinhydrone. In this range oxygen of the air does not affect the results. However, above pH = 8.5 oxygen reacts with the hydroquinone rather easily, and, further, in these alkaline solutions the ionization of the hydroquinone becomes sufficiently significant to affect the pH. Serious errors arise also in solutions containing appreciable quantities of salts and proteins. These salt and protein errors are due to the ionic strength effects of these on the activity coefficients of quinone and hydroquinone, and vitiate the assumption made in passing from equations (114) to (115) for the emf of the electrode.

The Glass Electrode. Fritz Haber and Z. Klemensiewicz [1] first showed that, when two solutions of different pH are separated by a glass membrane, a potential is established across the membrane whose magnitude depends on the difference in pH of the two solutions. Since then it has been found that when the pH of one of these solutions is held con-

[1] Haber and Klemensiewicz, Z. physik. Chem., **67**, 385 (1909).

stant while that of the second is varied, the emf of the *glass electrode* follows the equation

$$\mathcal{E}_G = \mathcal{E}_G^0 - \frac{RT}{\mathcal{F}} \ln a_{H^+}$$

$$= \mathcal{E}_G^0 + \left(\frac{2.3026\,RT}{\mathcal{F}}\right) pH \tag{121}$$

where $\mathcal{E}_G^0$ is a constant determined by the magnitude of the fixed pH and the particular arrangement employed. Equation (121) is strikingly similar to equation (115) for the quinhydrone electrode, and indicates that under proper conditions the glass electrode should be suitable for pH measurements.

A glass electrode assembly frequently utilized for pH work is illustrated in Fig. 9. In this figure A is the glass electrode, B is the solution
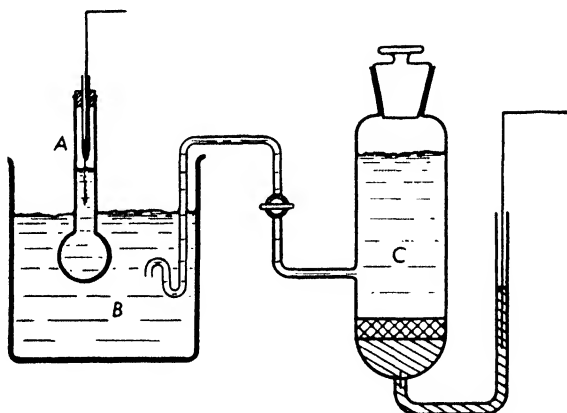
Fig. 9. Glass Electrode Assembly for pH

under test, and C is a calomel electrode used to complete the cell arrangement. The glass electrode, immersed directly in the solution whose pH is to be determined, consists of a thin bulb constructed usually of a special glass (Corning 015 glass composed of 72 per cent silica, 22 per cent sodium monoxide, 6 per cent calcium oxide). On the inside of this bulb is placed 0.1 N hydrochloric acid and a silver-silver chloride electrode. This whole combination, namely,

$$\text{Ag} \mid \text{AgCl(s)}, \ 0.1 \ \text{N HCl} \mid \text{glass} \mid$$
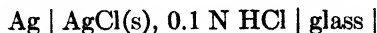
constitutes the glass electrode. When combined with the calomel the resulting cell is

$$\text{Ag} \mid \text{AgCl(s)}, \ 0.1 \ \text{N HCl} \mid \text{glass} \mid \text{Solution (pH} = x) \mid \text{Calomel} \tag{122}$$

from whose measured emf the single electrode potential of the glass electrode can be deduced, and therefrom the pH.

Because of the high internal resistance of the glass electrode, which may amount to as much as 100 million ohms, ordinary potentiometers cannot be used to measure the emf of cell (122). Either quadrant electrometers or vacuum tube voltmeters which require practically no current for their operation must be used. Since the advent of the glass electrode, various vacuum tube circuits have been developed which not only are sensitive to 0.01 pH unit or better, but also are portable and very rugged. These are readily available on the market.

Although $\mathcal{E}_G^0$ in equation (121) can be determined and recorded for any particular electrode assembly, it is found that such values of $\mathcal{E}_G^0$ are not applicable to all electrodes because of the "asymmetry potential" of the glass, which may be different for different electrodes. For this reason it is best to determine $\mathcal{E}_G^0$ for each electrode before use. To do this a solution of definite pH is placed in B, the emf of the cell is measured, and $\mathcal{E}_G^0$ is evaluated. Then the first solution is discarded, the solution under test substituted, the measurement repeated, and the pH calculated with the $\mathcal{E}_G^0$ found above. In this manner any "asymmetry potential" is incorporated into $\mathcal{E}_G^0$ of the electrode, and errors are avoided.

The solution most frequently used for calibration of glass electrode assemblies is 0.05 molar potassium acid phthalate. This solution has a pH of exactly 4.00 between 10° C and about 30° C, and a pH of 4.02 at 38° C.

The glass electrode is the nearest approach to a universal pH electrode known at present. It is not poisoned easily, nor is it affected by oxidizing or reducing agents or by organic compounds. Further, in various forms it can be used on quantities of solution as small as a fraction of a cubic centimeter. Its only limitations arise in strongly alkaline solutions where the glass is attacked, and in solutions of pH = 9 and above where presence of various cations, particularly sodium, leads to appreciable errors. Indications are also that below pH = 1 anions may cause deviations from strict applicability of equation (121). For these reasons glass electrodes of Corning 015 glass can be considered to be strictly satisfactory only between pH 1 and 9. With suitable corrections this range may be extended up to pH = 12. However, a new glass was developed recently with which good results can be obtained up to pH = 11 or 12 and, with small allowance for the sodium error, these may be extended up to pH = 13 or 14.

Application of pH Measurements. Besides the determination of pH *per se*, potentiometric pH measurements can be utilized in estimating the degree of ionization of weak acids and bases, and the degree

of hydrolysis of salts in the same manner as colorimetric methods described in the preceding chapter. From such data ionization and hydrolytic constants can readily be obtained. Potentiometric methods are now used much more extensively than colorimetric estimations of pH because they are capable of greater accuracy and, further, are applicable to a larger variety of substances.

Potentiometric Acid-Base Titrations. In Fig. 10(a) is shown a characteristic plot for the variation of pH with volume of added base
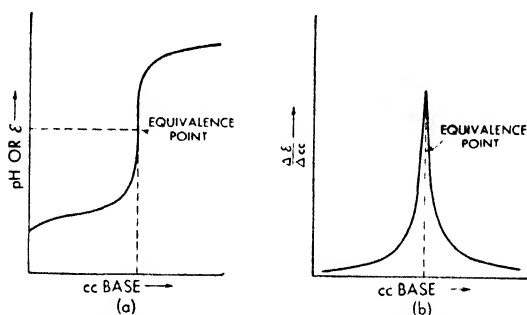
Fig. 10. Potentiometric Titration of an Acid with a Base

during the titration of an acid such as hydrochloric acid with a base like sodium hydroxide. It will be observed that the pH of the solution rises gradually at first, then more rapidly, until at the equivalence point there is a very sharp increase in pH for a very small quantity of added base. Once past the equivalence point, the curve again tapers off, indicating that the pH increases only slightly on the addition of excess base.

These changes in pH during a titration may be followed potentiometrically by immersing in the solution being titrated an electrode reversible to hydrogen ions, and coupling it with a suitable reference electrode. Since the potential of the latter remains constant, the emf of such a cell will vary only with the pH of the solution during the addition of base. Further, since the pH of any electrode reversible to hydrogen ions is proportional to the potential of the electrode, the cell emf will exhibit a course parallel to the curve shown in Fig. 10(a). Consequently, by measuring this emf at each stage of the titration and plotting it against the cubic centimeters of base, we may deduce from the plot the volume of base at which a sharp rise in emf occurs, and hence the equivalence point. This is the principle of all potentiometric acid-base titrations.

A more sensitive and satisfactory means of deducing the end point is to plot, not $\mathcal{E}$ of the cell against the volume, but the slope of the curve, i.e., $\Delta\mathcal{E}/\Delta cc$ vs. cc, Fig. 10(b). As the variation of the slope is greatest at the equivalence point, the latter plot exhibits a maximum at a volume corresponding to the end point of the titration. In practice a known

volume of the acid is placed in a beaker, the electrodes are inserted, and the base is added with stirring in rather large increments at the start. After each addition the emf is read. When the latter begins to show a tendency to increase appreciably, the increments are made smaller and smaller until near the end point only 0.1 cc or less are added at a time. By such a procedure it is possible to obtain a sharp maximum, and hence within narrow limits the volume of base equivalent to the acid taken.

The potentiometric method of deducing the acid or basic content of solutions possesses a number of distinct advantages over the ordinary methods of titration involving indicators. In the first place, indicators cannot be employed when solutions have a strong color of their own, whereas potentiometric titrations are not subject to his limitation. Second, indicators for any acid-base titration must be chosen so that the pH at which the indicator changes color corresponds more or less to the pH of the solution at the equivalence point. This means that some information is required *a priori* concerning the relative strengths of the acid and base involved. When these are both strong, little difficulty is encountered, and indicators such as methyl orange or phenolphthalein will both do. However, when this is not the case, the results obtained will be determined to a large degree by the good judgment exercised in choosing the correct indicator. This is not true of potentiometric titrations. These always yield the *equivalence point* whether this point comes exactly at the neutral point, or on the acid or basic side. Hence the uncertainty attending the use of indicators is removed.

Still a third advantage of the potentiometric method appears in the titration of polybasic acids or of mixtures of a strong and a weak acid with a base. With indicators it is as a rule impossible or difficult to titrate a polybasic acid in steps corresponding to different stages of neutralization. This is also the case with mixtures of strong and weak acids unless one indicator can be found whose color change corresponds to the neutralization of the strong acid, and another to that of the weak. Then, if the color of the first does not interfere with the second, the titration is conducted until the first indicator changes color, the second next added, and the titration completed. However, when the ionization constants of the different stages of a polybasic acid or of the different acids in a mixture differ by a factor of at least 1000, the potentiometric titration yields directly a number of distinct steps for the various neutralizations. This may be seen from Fig. 11, where the plot for the potentiometric titration of phosphoric acid with sodium hydroxide is shown. Point A indicates the neutralization of the first ionizable hydrogen, B of the second. The third cannot be thus obtained, because beyond B the solution is too alkaline for the sodium hydroxide to produce any signifi-

cant difference in pH. To get around this difficulty calcium chloride is added to the solution at a point such as C. The calcium chloride reacts
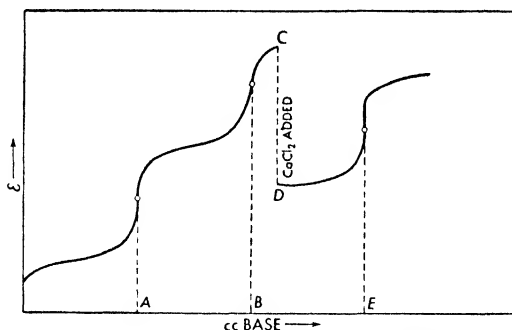
Fig. 11. Stepwise Potentiometric Titration of H_3PO_4 with NaOH

with the disodium phosphate present to precipitate calcium phosphate, and to produce a quantity of hydrochloric acid exactly equal to the acid content of the disodium phosphate according to the equation:

$$2 \, Na_2HPO_4 + 3 \, CaCl_2 = Ca_3(PO_4)_2 \downarrow + 4 \, NaCl + 2 \, HCl$$
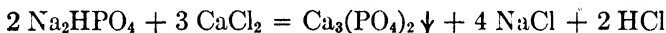
Because of the appearance of this strong acid the pH of the solution falls from C to a point D, making thus possible further titration from D onward until the third break E, corresponding to complete neutralization, results.

It is readily evident that by such a procedure it is possible to analyze mixtures of, say, phosphoric acid, sodium dihydrogen phosphate, and disodium phosphate. Likewise other combinations may be titrated in stages. But, it is of interest to point out that such stepwise neutralization cannot be obtained with sulfuric acid. In this acid partial neutralization of the first hydrogen leads to appreciable ionization of the second, and the two titrate as if sulfuric acid were a single monobasic acid.

Oxidation-Reduction and Precipitation Titrations. Oxidation-reduction titrations can be carried out potentiometrically in a manner identical to acid-base neutralizations by substituting for the electrode reversible to hydrogen ions an inert metal, such as a platinum wire. This metal acts as an oxidation-reduction electrode whose emf is determined by the activity ratio of the substance being oxidized or reduced. The variation of the emf of the cell with volume of reagent added follows essentially a curve similar to Fig. 10(a). As the sharp change in emf occurs at the equivalence point, these plots may again be utilized to detect the end points. In this way may be titrated various reducing agents with oxidants, or vice versa, without dependence on color change or oxidation-reduction indicators. Further, in certain instances several substances may be estimated in the same solution by a single titration.

The same type of behavior is observed also in many precipitation re-
actions when an electrode reversible to one of the ions involved is used.
Thus, in the titration of silver nitrate with sodium chloride a silver elec-
trode will exhibit a sharp change in emf when all the silver is precipitated.
Likewise a silver electrode will indicate the precipitation of all the chlo-
ride when the latter is titrated with silver nitrate. The power and speed
of the potentiometric method is particularly well illustrated in the esti-
mation of mixed halides. The analysis of solutions composed of chlorides,
bromides, and iodides by ordinary methods is long, involved, and tedi-
ous. Potentiometrically, however, such mixtures can be analyzed under
proper conditions by a single titration with silver nitrate in presence of
a silver electrode. The three "breaks" obtained give first the precipita-
tion of the least soluble silver iodide, next the intermediately soluble
silver bromide, and finally the precipitation of the most soluble silver
chloride. From the volumes of silver nitrate required to produce these
"breaks" the amount of each halide present in the mixture can readily
be calculated.

REFERENCES FOR FURTHER READING

See list at end of preceding chapter. Also:

1. Creighton and Koehler, "Principles and Applications of Electrochemistry,"
   third edition, John Wiley & Sons, Inc., New York, 1935, Vol. I.
2. M. Dole, "Experimental and Theoretical Electrochemistry," McGraw-Hill Book
   Company, Inc., New York, 1935.
3. M. Dole, "The Glass Electrode," John Wiley & Sons, Inc., New York, 1941.
4. S. Glasstone, "Introduction to Electrochemistry," D. Van Nostrand Company,
   Inc., New York, 1942.
5. R. W. Gurney, "Ions in Solution," Cambridge University Press, Cambridge,
   Eng., 1936.
6. Kolthoff and Furman, "Potentiometric Titrations," John Wiley & Sons, Inc.,
   New York, 1931.
7. W. M. Latimer, "The Oxidation States of the Elements and Their Potentials in
   Aqueous Solutions," Prentice-Hall, Inc., New York, 1938.
8. D. A. MacInnes, "The Principles of Electrochemistry," Reinhold Publishing
   Corporation, New York, 1939.

PROBLEMS

Note: In the following problems the temperatures are to be taken as 25° C unless
otherwise specified.

1. The resistance of a potentiometer for the full scale reading of 1.6 volts is 200
   ohms. What steady current must be maintained through the potentiometer
   coils so that the instrument may read volts directly?   *Ans.* 0.008 amp
2. Write the individual electrode reactions and the total cell reaction for each of
   the following cells. In each cell the negative electrode is at the left.

   (a) Pb | PbSO$_4$(s), SO$_4^{--}$ ‖ Cu^{++} | Cu
   (b) Cd | Cd^{++} ‖ H$^+$ | H$_2$(g)
   (c) Zn | Zn^{++} ‖ Fe^{+++}, Fe^{++} | Pt
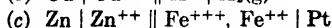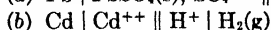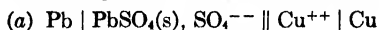
3. In each of the following cells write the cell reaction, and designate which reactions are spontaneous as written.

Sign of emf

(a) $Ag \mid AgCl(s), Cl^- \parallel I^-, AgI(s) \mid Ag$ $-$

(b) $K \mid K^+ \parallel Cl^-, Hg_2Cl_2(s) \mid Hg$ $+$

(c) $Pt \mid Tl^+, Tl^{+++} \parallel Cu^{++} \mid Cu$ $-$

(d) $Ni \mid Ni^{++} \parallel Pb^{++} \mid Pb$ $+$

4. A silver-silver ion electrode is measured against a normal calomel electrode at 25° C. The calomel electrode is negative, and the observed emf is 0.2360 volt. Write the cell reaction and calculate the potential of the silver-silver ion electrode.   *Ans.* $\mathcal{E} = 0.5160$ volt

5. Assuming that the substances present in each of the cells of problem 2 are in their standard states, calculate the emf of each cell and the free energy change accompanying the cell reaction.

6. At 25° C the free energy of formation of $H_2O(l)$ is $-56,560$ cal/mole. What will be the reversible emf at 25° C of the cell:

$$H_2(g, 1 \text{ atm}) \mid H^+ \parallel OH^- \mid O_2(g, 1 \text{ atm})$$
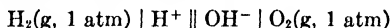
What is the single electrode potential of the $OH^- \mid O_2$ electrode?

7. The heat of formation of $H_2O(l)$ at 25° C is $-68,400$ cal/mole. Calculate the temperature coefficient of emf at 25° C of the cell given in problem 6.

8. From the equation giving the emf of the Weston cell as a function of temperature calculate ΔF, ΔH, and ΔS for the reaction occurring within the cell at 25° C.

9. Calculate the oxidation potentials of each of the following single electrodes:

(a) $Ag \mid AgCl(s), Cl^-(a = 0.0001)$

(b) $Pt \mid I_2(s), I^-(a = 1.5)$

(c) $Sn \mid Sn^{++}(a = 0.01)$

(d) $H_2(g, 1 \text{ atm}) \mid H^+(a = 15)$

10. Calculate the potentials, and specify the electrode polarity for the spontaneous reaction, for the cells obtained by combining (a)-(d), (b)-(d), and (a)-(c) in problem 9.

11. Calculate the potential of the cell:

$$Fe \mid Fe^{++}(a = 0.6) \parallel Cd^{++}(a = 0.001) \mid Cd$$
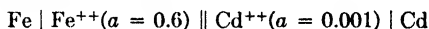
What must be the polarity of the Cd electrode in order for the cell to serve as a source of energy?   *Ans.* -0.043 volt; negative

12. Devise a cell in which the following reaction will take place:

$$\frac{1}{2} Br_2(l) + Fe^{++} = Br^- + Fe^{+++}$$
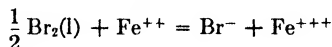
What will be the standard emf of the cell, and the standard free energy change in calories accompanying the reaction? Is the reaction as written spontaneous?

13. For the cell

$$Ag \mid AgBr(s), Br^-(a = 0.10) \parallel Cl^-(a = 0.01), AgCl(s) \mid Ag$$
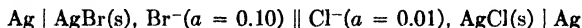
write the cell reaction, and calculate the potential at 25° C. Is the reaction as written spontaneous?

14. Calculate the potential at 25° C of the cell:

$$H_2(P = 1 \text{ atm}) \mid HBr(a_\pm = 0.2), Hg_2Br_2(s) \mid Hg$$
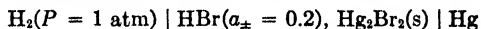
Ans. 0.2212 volt

15. The potential of the cell

$$Cd \mid CdI_2, AgI(s) \mid Ag$$
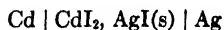
is 0.2860 volt at 25° C. Calculate the mean ionic activity of the ions in the solution, and the activity of the electrolyte.   *Ans.* $a_\pm = 0.392$; $a_{CdI_2} = 0.0603$

16. Indicate the cell in which the equilibrium constant for the following reaction can be measured:

$$\frac{1}{2} H_2(g, 1\ atm) + AgI(s) = H^+ + I^- + Ag(s)$$
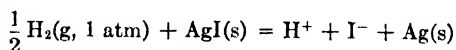
What is the equilibrium constant for this reaction at 25° C?
                                                      Ans. $K = 2.67 \times 10^{-3}$

17. From the results of the preceding problem, calculate the activity of HI in equilibrium with AgI(s), Ag(s), and H_2(g, 1 atm) at 25° C.

18. Write the cell reaction and calculate the potential of the cell:

$$H_2(g, 0.4\ atm)\ |\ HCl(a_\pm = 3.0),\ AgCl(s)\ |\ Ag$$
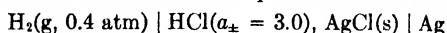
19. Finely divided Ni is added to a solution in which the molality of Sn^{++} is 0.1. What will be the activities of Ni^{++} and Sn^{++} when equilibrium is established? Assume that the molality is equal to activity for this calculation.
                              Ans. $a_{Ni^{++}} = 0.10$; $a_{Sn^{++}} = 6.0 \times 10^{-5}$

20. An excess of solid AgCl is added to a solution containing a concentration of Br^- of 0.1. Assuming activity equal to ionic concentration, calculate the concentration of Cl^- and Br^- ions at equilibrium.

21. Write the cell reaction and calculate the potential of the following cell:

$$Zn\ |\ Zn^{++}(a = 0.01)\ \|\ Fe^{++}(a = 0.001),\ Fe^{+++}(a = 0.1)\ |\ Pt$$
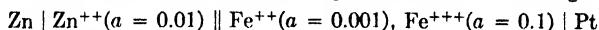
22. It is desired to determine the $\mathcal{E}^0$ value of the $Cu\ |\ Cu^{++}$ electrode and the mean ionic coefficients of $CuSO_4$. Explain what cells may be assembled, what measurements made, and derive an expression for the emf of the cells chosen. Assume that the $\mathcal{E}^0$ value of the reference electrode used is known.

23. Write the cell reaction and calculate the potential of the cell:

$$H_2(g, 1\ atm)\ |\ H_2SO_4(m = 0.05,\ \gamma_\pm = 0.397),\ Hg_2SO_4(s)\ |\ Hg$$
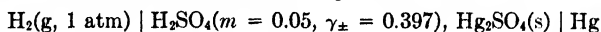
24. Explain what cell may be employed to measure the mean ionic activity coefficients of $CdCl_2$. Derive an expression for the emf of such a cell.

25. For the cells

$$Na(s)\ \left|\ \begin{matrix} NaI\ in \\ C_2H_5NH_2 \end{matrix}\ \right|\ Na(Hg)(0.206\%)$$

$$Na(Hg)(0.206\%)\ |\ NaCl(m = 1.022,\ \gamma_\pm = 0.650),\ Hg_2Cl_2(s)\ |\ Hg$$
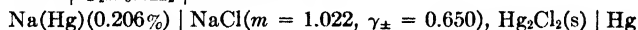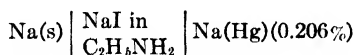
Lewis and Randall ("Thermodynamics," McGraw-Hill Book Company, Inc., New York, 1923, pp. 415–416) give potentials of 0.8453 and 2.1582 volts respectively. Write the reaction for each cell, and find from the data the $\mathcal{E}^0$ value for the sodium electrode.

26. The potential of the cell

$$Zn(s)\ |\ ZnCl_2(m = 0.01021),\ AgCl(s)\ |\ Ag$$
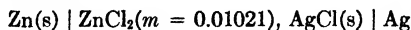
was found to be 1.1558 volts. What is the mean ionic activity coefficient of $ZnCl_2$ in this solution?

27. Keston [J. Am. Chem. Soc., **57**, 1671 (1935)] found that for cells of the type, $H_2(g, 1\ atm)\ |\ HBr(m),\ AgBr(s)\ |\ Ag(s)$ the emf's at various molalities of HBr are:

m	0.0003198	0.0004042	0.0008444	0.001355	0.001850	0.002396	0.003719
$\mathcal{E}$	0.48469	0.47381	0.43636	0.41243	0.39667	0.38383	0.36173

Obtain $\mathcal{E}^0$ for the cell by a graphical method, and then calculate the activity coefficients at each molality. Compare the activity coefficients thus obtained with those calculated from the Debye-Hückel limiting law.

28. The potential of the cell

$$Hg \mid Hg_2Cl_2(s), \ KCl \ (sat'd) \mid Cl_2(P = 0.283 \ atm)$$
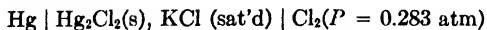
was observed to be 1.0746 volts. Using the $\mathcal{E}^0$ value for the calomel electrode given in the text, calculate the $\mathcal{E}^0$ value for the chlorine electrode.

29. For the cell

$$Zn \mid ZnSO_4, \ PbSO_4(s) \mid Pb(Hg)$$

in which the right-hand electrode was Pb saturated with Hg, and having an $\mathcal{E}^0$ value of 0.1207 for *oxidation*, the following data were obtained by Cowperthwaite and LaMer [J. Am. Chem. Soc., **53**, 4333 (1931)]:

Molality of ZnSO₄	Emf
0.0005	0.61144
0.002	0.58319
0.01	0.55353
0.05	0.52867

Calculate the mean ionic activity coefficient of $ZnSO_4$ in each of these solutions.

30. Write the cell reaction and calculate the potential of the cell

$$Cl_2(P = 0.9 \ atm) \mid NaCl \ (sol'n) \mid Cl_2(P = 0.1 \ atm)$$
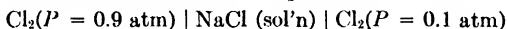
Will the cell reaction be spontaneous as written?

31. Given the cells:

   (a) $H_2(P = 1 \ atm) \mid HBr(a_{\pm} = 0.001), \ AgBr(s) \mid Ag$

   (b) same as (a) with $a_{\pm} = 2.5$

   (c) same as (a) with $a_{\pm} = 0.01$
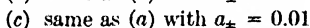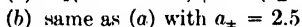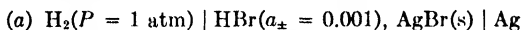
   show how cell (a) may be combined with either (b) or (c) to give a positive emf, and calculate the emf in each case.

32. What will be the cell reaction and the potential of the cell

$$Zn(s) \mid ZnCl_2(m = 0.02, \ \gamma_{\pm} = 0.642), \ AgCl(s) \mid Ag-Ag \mid AgCl(s),$$
$$ZnCl_2(m = 1.50, \ \gamma_{\pm} = 0.290) \mid Zn(s)$$
Ans. 0.1359 volt

33. Find the cell reaction and calculate the potential of the cell

$$Cd(s) \mid CdSO_4(m = 0.01, \ \gamma_{\pm} = 0.383), \ PbSO_4(s) \mid Pb-Pb \mid PbSO_4(s),$$
$$CdSO_4(m = 1.00, \ \gamma_{\pm} = 0.042) \mid Cd(s)$$
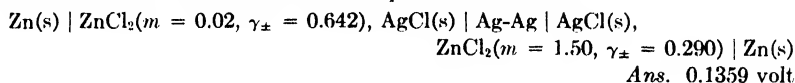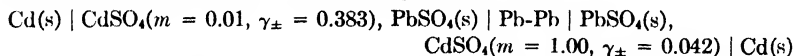
34. Find the cell reaction and calculate the potential of the following cell with transference:

$$Pb(s) \mid PbSO_4(s), \ CuSO_4(m = 0.2, \ \gamma_{\pm} = 0.110) \mid CuSO_4(m = 0.02,$$
$$\gamma_{\pm} = 0.320)PbSO_4(s) \mid Pb$$
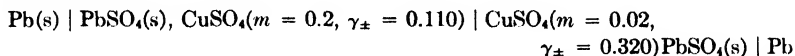
The transport number of Cu^{++} ion may be taken as 0.370. Is the cell reaction spontaneous as written?   *Ans.* 0.0412 volt; yes

35. Given the cell with transference:

$$H_2(P = 1 \ atm) \mid H_2SO_4(m = 0.005, \ \gamma_{\pm} = 0.643) \mid H_2SO_4(m = 2.00,$$
$$\gamma_{\pm} = 0.125) \mid H_2(P = 1 \ atm)$$
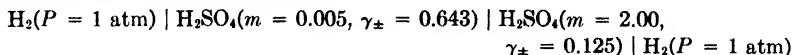
find the cell reaction and emf. The transport number of hydrogen ion is 0.821.

36. It is desired to determine the transport numbers of the ions in $CdCl_2$ by emf measurements. Show what cells would have to be assembled, and what measurements and calculations would have to be made.

37. The emf of the following cell:

$$Ag(s) \mid AgCl \ (sat'd), \ KCl(m = 0.05, \ \gamma_{\pm} = 0.840) \mid KNO_3 \mid AgNO_3(m = 0.1,$$
$$\gamma_{\pm} = 0.723) \mid Ag(s)$$
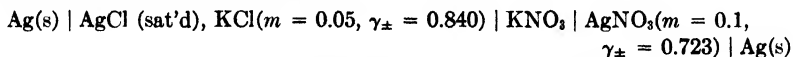
is 0.4312 volt. Estimate from these data the solubility product of AgCl.

Ans. 1.55×10^{-10}

38. Show how the proper $\mathcal{E}^0$ values from Table 3 may be combined and used to calculate the thermodynamically exact solubility product of $PbSO_4$.

Hint: Construct cell in which the reaction is

$$PbSO_4(s) = Pb^{++}(a = 1) + SO_4^{--}(a = 1)$$
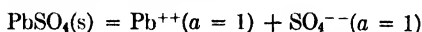
39. Using $\mathcal{E}^0$ values listed in Table 3, calculate $\mathcal{E}^0$ and ΔF^0 for the electrode process:

$$Fe \longrightarrow Fe^{+++} + 3 \ominus$$
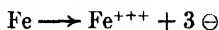
Ans. $\mathcal{E}^0 = 0.0448$ volt; $\Delta F^0 = -3100$ cal

40. The emf of the cell

$$H_2(g) \mid buffer \parallel Normal\ Calomel\ Electrode$$
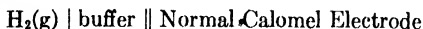
is 0.6885 volt at 40° C when the barometric pressure is 725 mm. What is the pH of the solution?   *Ans.* 3.44

41. The buffer in the cell of the preceding problem is replaced with a standard buffer of pH = 4.025, while all other conditions remain the same. What should then be the cell potential?

42. The potential of the cell

$$Pt \left| \begin{matrix} Buffer \\ Sat'd\ Quinhydrone \end{matrix} \right\| Normal\ Calomel\ Electrode$$
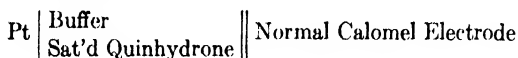
is 0.0042 volt at 35° C. What is the pH of the buffer?

43. The buffer used in the preceding problem is replaced by another buffer. The polarity of the cell is thereby reversed and a potential of 0.2175 volt observed. What is the pH of the second buffer?

44. Compare the emf's and the polarities which result when (a) a H_2 electrode ($P_{H_2} = 1$ atm) is placed in a buffer solution of pH = 6.55 using a normal calomel reference, and (b) the H_2 electrode is replaced by a quinhydrone electrode.

45. The cell

$$Ag(s) \mid AgCl(s),\ HCl(0.1\ N) \mid Glass \mid buffer \parallel Sat'd\ Calomel\ Electrode$$
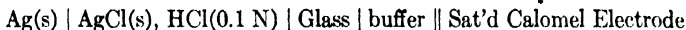
gives an emf of 0.1120 volt when the pH of the buffer used was 4.00. When a buffer of unknown pH was used, the potential was 0.3865 volt. What is the pH of the unknown buffer?

46. What would be the potential of the cell used in the preceding problem if a buffer of pH = 2.50 were used?

47. A 100-cc solution of 0.01 N HCl is titrated with 0.1 N NaOH solution. Calculate the pH of the solution after the addition of 0, 9.0, 9.9, 9.99, 10.0, 10.1, 11.0, and 20.0 cc of base. Plot the pH obtained against the volume of base added. Assume that the volume does not change during the course of the titration, and that the ionic concentration is equal to the ionic activity.

48. Repeat exercise 47 using 100 cc of 0.01 N CH_3COOH in place of the HCl. The dissociation constant of CH_3COOH may be taken as 1.8×10^{-5}.

49. One hundred cc of 0.01 N KCl is titrated with 0.1 N $AgNO_3$. Calculate the potential of a Ag electrode in the solution after the addition of 9, 10, and 11 cc of the reagent. Assume no volume change, and that ionic activity is equal to ionic concentration. The solubility product of AgCl may be taken as 1.7×10^{-10}.

Electrolysis and Polarization

Introduction. In the preceding chapter it was pointed out that an electrochemical cell may be employed to convert chemical energy into electrical and to transform electrical energy into chemical. We have seen that in performing the first of these functions the cell yields electrical energy as a result of a free energy decrease accompanying a spontaneous reaction within the cell. Under reversible conditions the electrical work thus obtained is exactly equal to the free energy decrease. When, on the other hand, the cell is operated irreversibly, the electrical energy available for doing work is less than the free energy decrease, and the difference becomes dissipated as heat.

Of necessity it must follow that any reaction that is nonspontaneous, i.e., one accompanied by a free energy increase, cannot be utilized for doing work. Instead, to make such reactions possible energy must be supplied to the system in order to raise the free energy content of the products above that of the reactants. Electrochemically this addition of energy is accomplished by passing electricity under a suitable applied potential through a cell in which the particular reaction can take place. By means of such *electrolysis* chemical changes are produced at the electrodes, and we obtain chemical action in nonspontaneous processes at the expense of the electrical energy.

From purely thermodynamic considerations it may be anticipated that the *minimum* electrical energy required to carry out a nonspontaneous reaction would be equal to the free energy increase accompanying the change; and this ΔF would in turn be equal, but be opposite in sign, to the free energy decrease attending the reverse spontaneous process. This would be true, of course, when the electrolysis is carried out reversibly, i.e., when the electrodes are completely reversible and when only a very minute current is passed through the cell. When these conditions of reversibility are not satisfied, the energy required will no longer be the theoretical minimum, but that plus the energy necessary to overcome the irreversibility. This means that in presence of any irreversibility the potential to be applied for electrolysis will have to be greater than the reversible emf of the cell. A cell requiring thus voltage

in excess of the theoretical is said to be *polarized*. The excess voltage is
called *polarization voltage*, while the phenomenon in general is referred
to as *polarization*.

Applied Potential and Electrolysis. In order to understand what
happens when an external voltage is applied to a cell, let us take first
the cell

$$Tl \mid Tl^+(a = 0.5) \parallel Sn^{++}(a = 0.01) \mid Sn$$
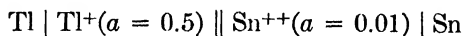
In this cell the thallium electrode is negative, and tends to send electrons
through the external circuit from the thallium to the tin. As a result of
this tendency the spontaneous reaction

$$Tl(s) + \frac{1}{2} Sn^{++}(a = 0.01) = \frac{1}{2} Sn(s) + Tl^+(a = 0.5) \tag{1}$$
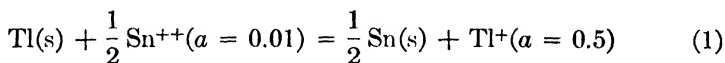
takes place, for which the potential is 0.1566 volt at 25° C. To reverse
this reaction an external potential must be applied to the cell of a mag-
nitude sufficient to overcome the emf of the cell itself, namely, 0.1566
volt. For this purpose an external source of emf must be connected to
the cell such that electrons enter the thallium electrode and leave through
the tin electrode; i.e., the negative pole of the external source of emf
must be connected to the negative electrode of the cell, the positive pole
to the positive electrode. Electrons will then flow externally from the
thallium to the tin as long as the applied potential is less than 0.1566 volt.
When the latter becomes equal to 0.1566 volt, the two emf's are balanced,
and no current flows in either direction. Finally, when the applied poten-
tial exceeds that of the cell, electrons from the external battery will enter
the thallium electrode, pass through the cell, and exit through the tin
electrode to complete the circuit. Since now the direction of current flow
is opposite to that of the spontaneous action of the cell, the reactions at
the electrodes are also reversed, and we obtain thus as a result of the
electrolysis the process opposite to that given in equation (1).

From this description it is evident that not all of the emf applied to
the cell is effective in passing current through the cell. Part of it, in this
case 0.1566 volt at 25° C, has to be utilized in overcoming the spon-
taneous emf of the cell itself. In general, if we let $\mathcal{E}$ be the applied
potential and $\mathcal{E}_b$ be the potential exerted by the cell itself, the net voltage
operative in current passage is $\mathcal{E} - \mathcal{E}_b$. With this net voltage the current
I resulting in a cell whose internal resistance is R follows according to
Ohm's law as

$$I = \frac{\mathcal{E} - \mathcal{E}_b}{R} \tag{2}$$

and this is the form this law takes in electrolysis when there is present
an opposing emf due to the cell. $\mathcal{E}_b$ is generally referred to as the *back*

emf of the cell in question. Under reversible conditions $\mathcal{E}_b$ is nothing more than the reversible emf of the cell. However, under irreversible conditions, i.e., when polarization is present, $\mathcal{E}_b$ is greater than the reversible emf, and a higher potential is required to yield a given current.

Ordinarily the cells encountered in electrolysis are not of the type discussed above. Rather they involve a *single electrolyte* into which dip a pair of electrodes. For instance, for the electrolysis of water the cell may consist of a pair of platinum electrodes immersed in a solution of an acid or base. Such a cell possesses no initial back emf, since both the electrodes and the electrolyte are the same. At first glance it may be expected, therefore, that in such a cell all of the external applied emf will go toward passage of electricity, with none of it necessary for overcoming back emf. However, this is not the case, for it is found experimentally that a potential of about 1.7 volts must be applied to the cell before there sets in a continuous evolution of hydrogen and oxygen. When a potential of less than 1.7 volts is applied, there is observed a momentary surge of current which rapidly falls to practically zero and stays there. The reason for this behavior is that, owing to the initial surge of current, small amounts of hydrogen are generated at the negative electrode and oxygen at the positive electrode. These gases adhere to the electrode surfaces, and convert these into hydrogen and oxygen electrodes with a back emf which operates to oppose the applied potential. At each potential below 1.7 volts the amounts of the two gases formed at the electrodes are just sufficient to make the back emf equal to the applied potential, and hence the first surge of current is rapidly reduced to practically zero. But, once the applied potential exceeds 1.7 volts, enough hydrogen and oxygen are formed to permit these gases to escape from the electrodes against the confining atmospheric pressure, and *continuous electrolysis* becomes possible. As the minimum voltage at which continuous evolution of hydrogen and oxygen sets in is about 1.7 volts, any voltage above this value is available for current passage. Therefore, by increasing the voltage above 1.7 volts higher and higher currents can be realized.

The significant fact in the above description is that *products of electrolysis convert even inert electrodes into active electrodes which can exercise a back emf.* This is true not only of gases but of other substances as well. Thus, as soon as some zinc is deposited on a piece of platinum, the latter is transformed into an active zinc electrode. Consequently, even where there is no initial back emf present, one is generated on attempted electrolysis by formation of a cell involving the products of the electrolysis.

Decomposition Potentials. Since even inert electrodes are converted on electrolysis into active electrodes with back emf, it becomes necessary to ascertain the minimum voltage required to produce continuous electrolysis of an electrolyte. This minimum voltage, called the

decomposition potential of an electrolyte, can be determined experimentally by means of the setup illustrated in Fig. 1. In this diagram A is the electrolytic cell containing the electrolyte to be studied, the two electrodes E and E', and the stirrer S. Connected to these electrodes is an external source of emf B, through a variable resistance C, and a milliammeter M for measuring the current passing through the circuit. By
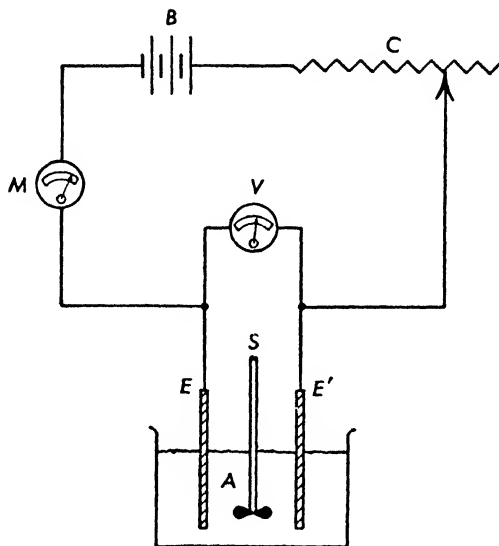
Fig. 1. Determination of Decomposition Potentials

means of the variable resistance C it is possible to regulate the voltage applied to the cell, and this voltage in turn can be measured by means of the voltmeter V inserted across the two electrodes. By using a high resistance voltmeter, all of the current indicated by the milliammeter may be considered to pass through the cell. Otherwise the reading of the milliammeter has to be corrected for the current passing through the voltmeter. The latter is given by the reading of the voltmeter divided by the internal resistance of the instrument.

To measure the decomposition voltage a series of current-applied voltage readings are taken both below and above the decomposition potential. At the start sufficient resistance is introduced at C to yield only a very small potential. After reading both the milliammeter and the voltmeter, some resistance is cut out at C, and another set of readings is taken. This procedure is continued until an appreciable current is passing through the circuit and electrolysis proceeds freely. During this operation it will be found that up to the decomposition voltage the current flowing will be either zero or very small even though the increase

in voltage is appreciable. On the other hand, once the decomposition voltage is exceeded, a sharp upturn in current for small increases in applied voltages takes place. If the current readings are plotted against the corresponding voltages, the curve resulting is of the type shown in Fig. 2. Along the branch AB, which is below the decomposition voltage, the current flowing is very small. Past B, however, the curve takes a
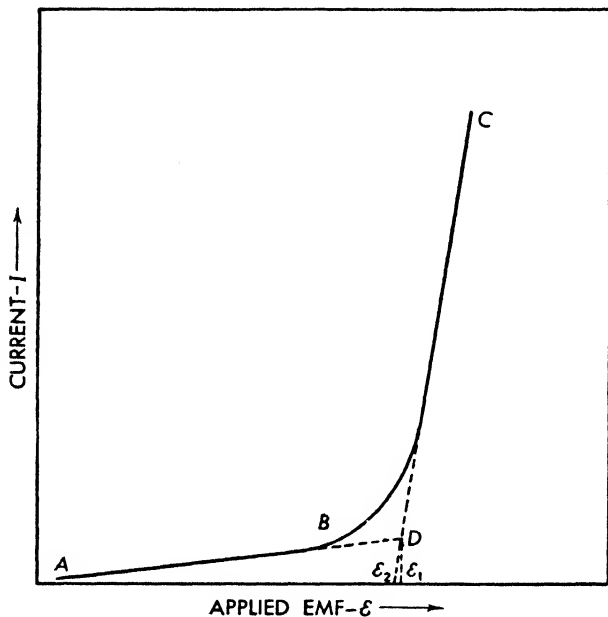
Fig. 2. Decomposition Potentials from I *vs.* ε Plots

sharp turn upward, indicating that a small change in voltage is effective in increasing the current appreciably. Farther along this branch, continuous electrolysis may be observed in the cell. Since the curve AB lies below the decomposition voltage while BC is above it, the voltage at which decomposition sets in must lie at the intersection of the two curves. To find this point the usual practice is to extend lines AB and BC until they intersect, as at D, and the voltage at this intersection, ε_1, is taken as the decomposition voltage. An alternate procedure is to extend line BC until it intersects the $I = 0$ axis, and the value of ε at the latter point, ε_2, is chosen as the voltage at which continuous electrolysis becomes possible. The idea here is that, since Ohm's law is followed more or less closely by the branch BC, extension of this law to $I = 0$ should yield the decomposition voltage even though there is some current flow before this voltage is attained. Despite the fact that the second method yields somewhat lower results than the first, there is no way of deciding

between them. Consequently either is chosen according to preference, with possibly the first more frequently than the second.

The appearance of current below the decomposition potential arises from the fact that the products of electrolysis deposited on the electrodes do not remain there, but tend to diffuse slowly back into the solution. To compensate for this loss by diffusion a small current must pass through the cell in order to regenerate the material lost. As the diffusion proceeds continuously, the current must be supplied continuously, and this accounts for the small but steady current observed below the decomposition potential.

TABLE 1

DECOMPOSITION POTENTIALS IN 1 N SOLUTIONS (Pt ELECTRODES)

Electrolyte	Dec. Potential E_d (Volts)	Products of Electrolysis	Rev. Decomp. Potential E_r	$E_d - E_r$
HNO_3	-1.69	$H_2 + O_2$	-1.23	-0.46
$CH_2ClCOOH$	-1.72	$H_2 + O_2$	-1.23	-0.49
$CHCl_2COOH$	-1.66	$H_2 + O_2$	-1.23	-0.43
$HClO_4$	-1.65	$H_2 + O_2$	-1.23	-0.42
H_2SO_4	-1.67	$H_2 + O_2$	-1.23	-0.44
H_3PO_4	-1.70	$H_2 + O_2$	-1.23	-0.47
$NaOH$	-1.69	$H_2 + O_2$	-1.23	-0.46
KOH	-1.67	$H_2 + O_2$	-1.23	-0.44
NH_4OH	-1.74	$H_2 + O_2$	-1.23	-0.51
$N(CH_3)_4OH$	-1.74	$H_2 + O_2$	-1.23	-0.51
HCl	-1.31	$H_2 + Cl_2$	-1.37	$+0.06$
HBr	-0.94	$H_2 + Br_2$	-1.08	$+0.14$
HI	-0.52	$H_2 + I_2$	-0.55	$+0.03$
$Cd(NO_3)_2$	-1.98	$Cd + O_2$	-1.25	-0.73
$CdSO_4$	-2.03	$Cd + O_2$	-1.26	-0.77
$CoCl_2$	-1.78	$Co + Cl_2$	-1.69	-0.09
$CoSO_4$	-1.92	$Co + O_2$	-1.14	-0.78
$CuSO_4$	-1.49	$Cu + O_2$	-0.51	-0.98
$Pb(NO_3)_2$	-1.52	$Pb + O_2$	-0.96	-0.56
$NiCl_2$	-1.85	$Ni + Cl_2$	-1.64	-0.21
$NiSO_4$	-2.09	$Ni + O_2$	-1.10	-0.99
$AgNO_3$	-0.70	$Ag + O_2$	-0.04	-0.66
$ZnBr_2$	-1.80	$Zn + Br_2$	-1.87	$+0.07$
$ZnSO_4$	-2.55	$Zn + O_2$	-1.60	-0.95

In column 2 of Table 1 are given the decomposition potentials obtained by the method described above in 1 N solutions of various electrolytes at room temperature. In all cases the electrodes were platinum. In each instance the products of electrolysis are given, as well as the reversible decomposition potentials calculated for the various electrolytes. The

manner of arriving at the latter and the significance of the observed results will now be explained.

Calculation of Reversible Decomposition Potentials. In order to interpret the observed decomposition potentials, the reversible decomposition potentials of the various electrolytes are required. Under reversible conditions these potentials would be the sum of the reversible emf's established at the electrodes by the products of electrolysis deposited upon them and the concentrations of the electrolyte with which they are in contact. Consequently, as soon as the products of electrolysis, the concentration of the electrolyte, and the direction of the electrode reaction, i.e., oxidation or reduction, are known, the reversible decomposition potentials can be calculated by the methods of the preceding chapter.

The direction of any electrode reaction in an electrolytic cell is readily ascertained from the polarity of the electrodes. In any electrolytic cell electrons enter through the electrode connected to the *negative* side of the external source of emf. At this electrode, therefore, called the *cathode*, the reaction must always be a *reduction*. On the other hand, electrons leave the cell through the electrode connected to the positive side of the external battery, and hence at this electrode, called the *anode* of the electrolytic cell, the reaction must always be an *oxidation*. With this information we are ready to consider the calculation of the reversible decomposition potentials of various electrolytes.

For this purpose let us take first the electrolysis of 1 N hydrochloric acid between platinum electrodes. This electrolysis involves the reduction of hydrogen ions to hydrogen at the cathode, and the oxidation of chloride ions to chlorine at the anode. Since at the decomposition potential the hydrogen and chlorine gases escape at the electrodes at 1 atmosphere pressure, we have for the two electrode reactions

$$2 H^+(N = 1) = H_2(g, 1 \text{ atm})$$
$$2 Cl^-(N = 1) = Cl_2(g, 1 \text{ atm})$$
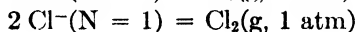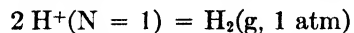
For the first of these electrodes the reversible emf follows from the Nernst electrode equation as

$$\mathcal{E}_{H_2} = \mathcal{E}^0_{H_2} - \frac{.059}{2} \log_{10} \left(\frac{1}{C_{H^+} f_{H^+}}\right)^2$$
$$= + .059 \log_{10} (0.811)$$
$$= - 0.005 \text{ volt}$$

Again, the emf of the chlorine electrode follows as

$$\mathcal{E}_{Cl_2} = \mathcal{E}^0_{Cl_2} - \frac{.059}{2} \log_{10} \left(\frac{1}{C_{Cl^-} f_{Cl^-}}\right)^2$$
$$= - 1.358 + .059 \log_{10} (0.811)$$
$$= - 1.363 \text{ volts}$$

Adding the two potentials, the reversible decomposition voltage of 1 N hydrochloric acid is, at 25° C,

$$\mathcal{E} = \mathcal{E}_{H_2} + \mathcal{E}_{Cl_2}$$
$$= -0.005 - 1.363$$
$$= -1.37 \text{ volts}$$

As a second example may be taken the electrolysis of zinc bromide. In this electrolyte the deposition of zinc on the cathode converts this electrode into a zinc electrode at which the reaction

$$Zn^{++} + 2 \ominus = Zn(s)$$
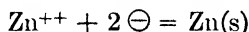
takes place, while the oxidation of bromide ions to bromine at the anode converts the latter electrode into a bromine-bromide electrode at which the process

$$2 Br^- = Br_2(l) + 2 \ominus$$

occurs. At 25° C, and in 1 N solution, the mean activity coefficient is $f = 0.38$, while the concentrations of the two ions are $C_{Zn^{++}} = 0.5$ molar and $C_{Br^-} = 1.0$ molar. We obtain, therefore, for the emf's of the two electrodes,

$$\mathcal{E}_{Zn} = \mathcal{E}^0_{Zn} - \frac{.059}{2} \log_{10} \frac{1}{C_{Zn^{++}} f_{Zn^{++}}}$$
$$= -0.761 + \frac{.059}{2} \log_{10} (0.5 \times 0.38)$$
$$= -0.782$$
$$\mathcal{E}_{Br_2} = \mathcal{E}^0_{Br_2} - \frac{.059}{2} \log_{10} \frac{1}{(C_{Br^-} f_{Br^-})^2}$$
$$= -1.066 + .059 \log_{10} (1 \times 0.38)$$
$$= -1.091$$

and, for the reversible decomposition potential of zinc bromide,

$$\mathcal{E} = \mathcal{E}_{Zn} + \mathcal{E}_{Br_2}$$
$$= -0.782 - 1.091$$
$$= -1.87 \text{ volts}$$

In order to calculate the reversible decomposition potentials in electrolytes where oxygen is evolved, the $\mathcal{E}^0$ value of the oxygen-hydroxyl ion electrode is required. Since the oxygen electrode is highly irreversible, this quantity cannot be obtained directly from emf measurements, but must be calculated from free energy data from other sources. For this purpose we may utilize the free energy changes in the following two reactions, namely,

$$H_2(g) + \frac{1}{2} O_2(g) = H_2O(l) \qquad \Delta F^0_{25° C} = -56,700 \text{ cal} \quad (3)$$

$$2 H_2O(l) = 2 H^+ + 2 OH^- \qquad \Delta F^0_{25° C} = +38,100 \text{ cal} \quad (4)$$

On adding these two equations, we obtain

$$H_2(g) + \frac{1}{2} O_2(g) + H_2O(l) = 2\,H^+ + 2\,OH^-$$

$$\Delta F^0 = -\ 18{,}600 \ \text{cal} \tag{5}$$
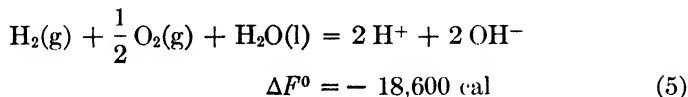
As 2 faradays are required to accomplish the reaction in equation (5), the $\mathcal{E}^0$ corresponding to the ΔF^0 of this process is

$$\Delta F^0 = -\ n\mathcal{F}\mathcal{E}^0$$
$$-\ 18{,}600 \times 4.185 = -2 \times 96{,}500 \ \mathcal{E}^0$$
$$\mathcal{E}^0 = 0.403 \ \text{volt}$$

However, for the reaction $H_2(g) = 2\,H^+$, $\mathcal{E}^0 = 0$. Subtracting this reaction from equation (5), we arrive then at the standard electrode potential for the solution of oxygen to form hydroxyl ions, namely,

$$\frac{1}{2} O_2(g) + H_2O(l) + 2 \ominus = 2\,OH^- \qquad \mathcal{E}^0_{25^\circ C} = 0.403 \ \text{volt} \tag{6}$$
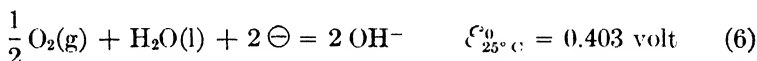
For the converse process, i.e., the deposition of oxygen at 1 atmosphere pressure from a solution containing hydroxyl ions at unit activity, the standard potential of the oxygen electrode follows, therefore, as

$$2\,OH^- = \frac{1}{2} O_2(g) + H_2O(l) + 2 \ominus \qquad \mathcal{E}^0_{25^\circ C} = -\ 0.403 \ \text{volt} \tag{7}$$
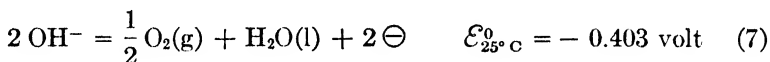
With the $\mathcal{E}^0$ of the oxygen electrode thus available, the reversible decomposition potentials for electrolytes involving evolution of oxygen on electrolysis may be calculated. As an initial example of this type may be taken the electrolysis of 1 N silver nitrate, which yields silver at the cathode and oxygen at the anode. For the reduction at the silver electrode thus formed at the cathode, we have the reaction $Ag^+(N = 1) = Ag(s)$. Hence the corresponding emf is

$$\mathcal{E}_{Ag} = \mathcal{E}^0_{Ag} - \frac{.059}{1} \log_{10} \frac{1}{C_{Ag^+} f_{Ag^+}}$$
$$= 0.799 + .059 \log_{10} (1 \times 0.396)$$
$$= 0.775 \ \text{volt}$$

Again, for the oxidation at the oxygen electrode formed at the anode, we have the process

$$2\,OH^- = \frac{1}{2} O_2(g) + H_2O(l)$$
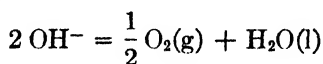
where, since the solution is essentially neutral, the concentration of hydroxyl ions may be taken as 10^{-7} gram ionic weights per liter. The potential of the oxygen electrode is, then,

$$
\begin{aligned}
\mathcal{E}_{O_2} &= \mathcal{E}_{O_2}^0 - \frac{.059}{2} \log_{10} \frac{1}{a_{OH^-}^2} \\
&= -0.403 + .059 \log_{10} (10^{-7}) \\
&= -0.816 \text{ volt}
\end{aligned}
$$

Adding the potentials of the silver and oxygen electrodes, the reversible decomposition voltage of the silver nitrate follows as

$$
\begin{aligned}
\mathcal{E} &= \mathcal{E}_{Ag} + \mathcal{E}_{O_2} \\
&= 0.775 - 0.816 \\
&= -0.04 \text{ volt}
\end{aligned}
$$

These examples show that the reversible decomposition potentials depend on the concentration of the electrolyte. However, this is not true when water is electrolyzed to yield hydrogen and oxygen, as in the electrolysis of many acids and bases. In the latter instance the overall cell reaction is given by the converse of equation (5), for which the emf follows as

$$
\begin{aligned}
\mathcal{E} &= \mathcal{E}^0 - \frac{.059}{2} \log_{10} \frac{1}{a_{H^+}^2 a_{OH^-}^2} \\
&= -0.403 + \frac{.059}{1} \log_{10} (a_{H^+} a_{OH^-})
\end{aligned}
$$

But, *in any aqueous solution* the product $a_{H^+} a_{OH^-}$ is a constant equal to 10^{-14} at 25° C. Hence,

$$
\begin{aligned}
\mathcal{E} &= -0.403 + .059 \log_{10} (10^{-14}) \\
&= -0.403 - 0.826 \\
&= -1.229 \text{ volts}
\end{aligned}
$$

and this should be, therefore, the reversible decomposition potential of any cell in which the essential process is the electrolysis of water. Further, this potential should be independent of the concentrations of the electrolyte or its nature. This deduction accounts for the fact that the reversible decomposition potentials in Table 1 for all the acids and bases yielding oxygen and hydrogen are the same.

Significance of Observed Decomposition Potentials. With this explanation of the calculation of reversible decomposition potentials of electrolytes, we are in a position to consider the significance of the experimental results given in Table 1. First of all, a comparison of the potentials given in columns 2 and 4 of the table for electrolytes that do not yield oxygen on electrolysis of 1 N solutions, namely, hydrochloric

acid, hydrobromic acid, hydriodic acid, cobalt chloride, nickel chloride, and zinc bromide, reveals that for these the observed decomposition potentials are not far different from those calculated theoretically. From this it may be concluded, therefore, that the ions in these substances, i.e., hydrogen, the metals, and the halogens, are deposited on platinum electrodes under conditions which approximate more or less to reversibility. On the other hand, in electrolytes in which oxygen is evolved, as in the acids and bases at the top of the table and in the salts cadmium nitrate, cadmium sulfate, cobalt sulfate, copper sulfate, lead nitrate, nickel sulfate, silver nitrate, and zinc sulfate, considerable discrepancy exists between observed and calculated decomposition potentials. In all these instances the potential actually required for decomposition is considerably greater than that calculated on the assumption of reversibility. Moreover, since, as has been pointed out, the metals and hydrogen are deposited essentially reversibly on platinum, this discrepancy indicates that the irreversibility observed must occur primarily in the evolution of oxygen at the platinum surfaces. In other words, oxygen electrodes, even at platinum surfaces, are appreciably polarized. In the case of the acids and bases the extent of this polarization, which for the moment may be ascribed totally to the oxygen, amounts to about 0.45 volt. However, in the deposition of oxygen from salt solutions the polarization is greater, and may be as much as a volt.

Another interesting point in Table 1 is the fact that acids and bases where oxygen and hydrogen are evolved have essentially the same decomposition potentials, namely, about − 1.7 volts. This observation is in line with the conclusion arrived at in the preceding section for the decomposition potential of any electrolyte in which the essential process is the electrolysis of water. However, the decomposition potential is not 1.23 volts as predicted theoretically, but that plus the polarization voltage of about 0.45 volt for the oxygen evolution, which is very nearly the same for all the acids and bases. In this connection it may be pointed out that whereas the decomposition potential of 1 N hydrochloric is − 1.31 volts, the potential increases as the acid is diluted, eventually reaching a value of − 1.7 volts. From this it may be concluded that, whereas in the more concentrated solutions the products of electrolysis are hydrogen and chlorine, in the more dilute solutions they are hydrogen and oxygen. This is actually the case. Between the two extremes both chlorine and oxygen are evolved at the anode in varying proportions, giving this electrode a potential intermediate between that of the oxygen and chlorine electrodes.

Types of Polarization. Measurements of decomposition potentials are not suitable for extended polarization studies. First, such data yield information on the sum of the polarization voltages at the two electrodes,

rather than on that at any one. Again, the results of decomposition potential measurements apply only to the conditions obtaining when continuous electrolysis sets in. They tell us nothing about what happens at higher current densities. Consequently, in order to ascertain the extent of polarization at each electrode separately and at various current densities, more suitable methods are required.

However, before turning to these it is necessary to point out that any observed polarization can be of two types, namely, (1) *concentration polarization*, and (2) *overvoltage*. The first arises from the fact that, with the changes in the concentration of the electrolyte about the electrodes during electrolysis, cells may arise which can develop an emf opposite to that applied. On the other hand, overvoltage is a polarization potential whose source lies in some process at the electrode which takes place irreversibly, and is thus a phenomenon intimately associated with the nature of the electrode and the processes occurring at its surface. Under ordinary conditions the total polarization observed will be the sum of both effects. But since the overvoltage is of particular electrochemical interest, it becomes important to understand the nature of concentration polarization, and the methods by which it may be minimized or eliminated in order to arrive at the overvoltage.

Concentration Polarization. The source of concentration polarization in an electrolytic cell may be illustrated with the following example. Suppose that a solution of 0.1 N silver nitrate is to be electrolyzed between *silver* electrodes. At the start the concentration of the electrolyte is uniform through the solution, the electrodes are identical, and so the cell has no initial back emf. However, as soon as current is sent through the cell, silver begins to dissolve at the anode while silver ions begin to deposit at the cathode. The solution in the immediate vicinity of the anode consequently becomes more concentrated in silver ions than 0.1 normal, that about the cathode less concentrated. As we have now two different concentrations about the electrodes, the combination must constitute a concentration cell whose emf is given by

$$\mathcal{E} = -\frac{RT}{n\mathcal{F}} \ln \frac{C_2 f_2}{C_1 f_1}$$

A moment's reflection on the nature of the concentration cell thus formed by the flow of current will reveal that the cell tends to send a current opposite to that of the external source, and hence that its emf operates against the applied potential. The result is thus a polarization requiring that part of the applied potential be utilized to overcome the emf developed due to the concentration changes about the electrodes.

Similarly it can be shown with other examples that in other types of cells as well current passage results in concentration changes about the

electrodes which tend to establish polarization potentials in the electrodes. Unless the concentration changes are permitted to become very large, these polarizations are at best rather small and can hardly account for the large polarization potentials observed in many cells. Further, some forces operating in solution, such as electrolytic conduction by the current and thermal diffusion, tend to equalize the concentration differences. As increase in temperature favors electrolytic conduction and diffusion, this equalization can be promoted by raising the temperature. However, a much more effective method of overcoming concentration polarization is efficient stirring of the electrolyte. In electrolytic cells where the electrolyte is thoroughly agitated the concentration polarization can be practically eliminated, and hence any polarization voltage observed under such conditions must be essentially overvoltage due to irreversibility in the electrode processes.
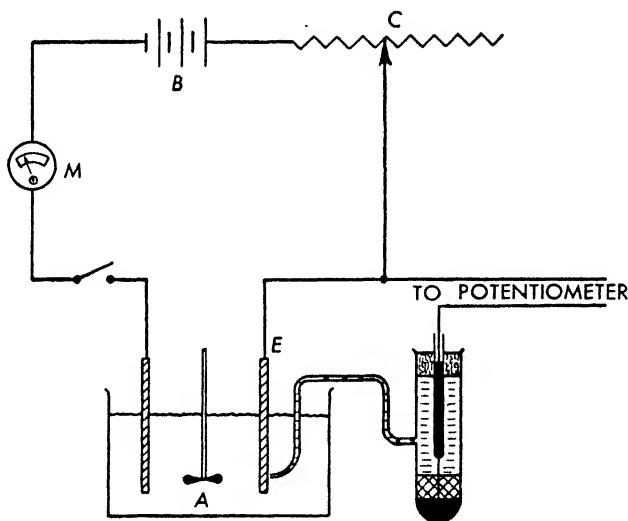
Fig. 3. Determination of Overvoltage

Measurement of Overvoltage. Direct measurements of overvoltage attending the deposition of electrolytic products at various electrodes can be made by observing the potentials of the electrodes when current is passing, and under conditions where there is no concentration polarization, i.e., in stirred solutions. For this purpose the setup generally employed is shown in Fig. 3. The electrolytic cell A, containing electrodes of known cross-sectional area, the electrolyte, and a stirrer, is connected in series with an external battery B, through a variable resistance C, an ammeter M, and a switch. Any desired current strength thus can be sent through the electrolyzing circuit by varying the resist-

ance at C. To obtain now the potential of, say, electrode E, the latter is made one-half of a new cell, the other half of which is a suitable reference electrode, such as a calomel, and the two are connected to a potentiometer. In order to avoid inclusion of any appreciable IR drop through the solution in measuring the potential of E against the reference electrode, the arm of the latter is usually drawn out to a fine tip and placed flush against the electrode.

To make a measurement the stirrer is started, the switch closed, and the current through the electrolyzing circuit, as indicated by M, is adjusted to any desired value by means of C. After the current has become steady, the emf of E against the reference electrode is read on the potentiometer, as well as the current given by M. The latter is then readjusted to some other value, and the whole operation is repeated. In this manner a series of emf readings can be obtained for various values of the current passing through the cell. On dividing the current strength at each point by the area of electrode E, the current density follows, while the emf of E at each stage is obtained by subtracting from the measured emf the potential of the reference electrode. Finally, in order to arrive at the overvoltage the reversible potential of E must be subtracted from the observed electrode potentials. These reversible emf's may be calculated, or they can be measured by taking the potential of E against a reference electrode when no current is flowing through the cell. In the latter instance an electrode material must be used at which the particular electrode reaction is reversible.

An alternate procedure is to permit the electrolysis in the cell to proceed until the back emf is established, and then to open the switch. When this is done it is found that the combination of E vs. the reference electrode acts for a brief period as a primary cell whose emf is high at first and then falls rapidly. By reading this back emf on the potentiometer at various time intervals after opening the electrolyzing circuit, a curve of voltage against time can be constructed which, on extrapolation back to the time at which the current was interrupted, yields the initial back emf. Since the emf usually falls rapidly, particularly at the higher current densities, this simple mode of operation is not very satisfactory. To obviate this difficulty commutating switches are frequently employed which turn the current on and off. With this arrangement the back emf is much more stable, and can be measured more conveniently. Another refinement is to use recording potentiometers on which the potentials are automatically plotted during the on and off cycles.

Overvoltage in Cathodic Deposition of Metals. Overvoltage studies on the cathodic deposition of metals indicate that generally speaking the deposition of metals takes place with relatively small over-

voltage. In line with theoretical requirements it is found that metals do not begin to plate on the cathode until the potential of the given metal·in the particular solution is exceeded. Thereafter the potential of the metal rises slowly with current density, indicating some polarization, but the overvoltage hardly ever exceeds several tenths of a volt at the higher current densities. Metals such as zinc, copper, and cadmium exhibit particularly low overvoltage, while metals of the iron group, namely, iron, cobalt, and nickel, show higher overvoltages. Just what is the cause of metal overvoltage is not yet clear. Various theories have been proposed, but at present too little information is available to make a choice among them.

The overvoltage of a particular metal is not the same for all salt solutions of the metal, but varies slightly with the nature of the co-ion. It is affected also by the presence of acids, colloids, and other added agents. Finally, the overvoltage of metals is increased when these are deposited from solutions in which the particular metal can exist as a complex salt, as silver in cyanide solutions.

Hydrogen Overvoltage. Of particular importance is the overvoltage exhibited by hydrogen in its deposition on various cathodic surfaces. *On platinized platinum and at zero current density the overvoltage of hydrogen is zero*, since these are the conditions under which the hydrogen electrode is reversible. However, as the current density is increased the overvoltage also goes up, but at no time does it become particularly large. In fact, at a current density of 1.5 amp per square centimeter it is only 0.05 volt at 25° C. On the other hand, at other cathodic surfaces, including smooth and shiny platinum, the overvoltage is much higher. The manner in which the overvoltage of hydrogen varies with the current density on several cathodic surfaces is shown in Fig. 4. These data, obtained in 2 N sulfuric acid at 25° C,[1] indicate that the overvoltage of hydrogen is low only on platinized platinum. On smooth platinum surfaces the overvoltage is low at very small current densities, but increases gradually with the latter to approach a limiting value at high current

Fog. 4. Hydrogen Overvoltage on Various Metals (2 N H_2SO_4 at 25° C)

---

[1] Knobel, Caplan, and Eiseman, Trans. Am. Electrochem. Soc., **43,** 55 (1923).

densities. This is true also of gold, graphite, tellurium, and palladium cathodes. However, on silver, mercury, bismuth, cadmium, tin, iron, zinc, nickel, and lead surfaces the hydrogen overvoltage is already quite large at zero current density. These electrodes, therefore, cannot function reversibly under any conditions. Hence it is not possible to deposit hydrogen on them at anywhere near the reversible potentials. Another characteristic of these metals with high hydrogen overvoltage is that they exhibit at low current densities a very rapid increase in overvoltage to values which remain essentially constant at the higher densities.

Frequently the variation of hydrogen overvoltage with current density at constant temperature can be represented by the equation

$$\mathcal{E}_o = a + b \log_{10} I \tag{8}$$

where $\mathcal{E}_o$ is the overvoltage, I the current density, and a and b are constants. According to this equation a plot of $\mathcal{E}_o$ against $\log_{10} I$ should be linear with slope equal to b and a y intercept of a. This expectation has actually been realized in a number of instances and over a fairly wide range of current densities. Furthermore, it has been found that many metals have the same slope b, equal to approximately 0.12. This value of the constant corresponds very closely to $(2 \times 2.303 \, RT)/\mathcal{F}$, which at 25° C is equal to 0.118. Further evidence for this association of b with $2 \times 2.303 \, RT/\mathcal{F}$ is the fact that study of the effect of temperature on overvoltage has shown that the slope of the overvoltage-$\log_{10} I$ curve increases linearly with the absolute temperature in accord with the relation

$$\mathcal{E}_o = a + \frac{(2 \times 2.303 \, RT)}{\mathcal{F}} \log_{10} I \tag{9}$$

Equation (9) does not hold invariably, however. Although in many instances this equation is approximated fairly closely by the data, in certain others the discrepancy may be quite large. Thus slopes have been obtained as low as 0.025 on platinized platinum, and as high as 0.3 on lead and graphite cathodes. Furthermore, equations (8) or (9) are insufficient frequently to represent the variation of overvoltage with current density over the full range of the latter. Another complicating factor encountered occasionally is a variation of overvoltage with time, indicating that changes are taking place in the nature of the electrode surface.

Besides the nature of the cathode material and the current density, several other factors greatly influence hydrogen overvoltage. Increase of temperature has been found to decrease the overvoltage, although, as we have seen, the slope of the overvoltage-current density curve is steeper at higher temperatures. Thus, for example, the overvoltage on

copper at 0.4 amp per square centimeter decreases from 0.57 volt at 0° to 0.40 volt at 74° C. On mercury and several other electrodes the decrease in hydrogen overvoltage with temperature amounts to about 2 millivolts per degree temperature rise. Measurements of hydrogen overvoltage at various pressures, both above and below atmospheric, indicate that at higher pressures the overvoltage is changed only slightly, but at the lower pressures the overvoltage rises sharply on copper, nickel, and mercury cathodes. Overvoltage is also influenced by the presence of various impurities and additions to the cathode materials.

The meager data available seem to show that the composition and pH of the solution exert little effect on the hydrogen overvoltage. On the other hand, the nature of the cathode surface is very important. On smooth, shiny, and polished surfaces the overvoltage is invariably greater than on rough, pitted, or etched surfaces. This difference is particularly well illustrated with the hydrogen overvoltage on smooth and platinized platinum. Analogous differences, though not so large, have been observed also on other metals.

Theories of Hydrogen Overvoltage. When the various stages involved in any overall electrode reaction proceed rapidly, equilibrium can readily be established between the ions in solution and the electrode, and the latter can thus behave reversibly and exhibit its equilibrium potential. On the other hand, if any of the intermediate processes for one reason or another are slow, or if any complicating reactions take place between the electrode and some of the substances appearing in the electrode reaction, equilibrium cannot be established, and hence the electrode cannot behave reversibly. In seeking an explanation of hydrogen overvoltage, therefore, the various stages in the formation of hydrogen gas from hydrogen ions in solution must be considered from the standpoint of speed and attainment of equilibrium. Again, the possibility of reaction between hydrogen atoms or molecules and the electrode must also be scrutinized.

Of the many theories of hydrogen overvoltage proposed, that of Newbery [1] ascribes hydrogen overvoltage to formation of hydrides between the electrode and the atomic hydrogen liberated on the electrodes. The hydride film formed on the surface is believed then to offer a high "transfer resistance" to current passage, with the result that an appreciable potential drop occurs, and becomes included in the measured electrode potential as an overvoltage. However, although hydride formation has actually been observed in some instances, and although transfer resistance may possibly be a factor, it is highly questionable whether these theories are generally applicable and can serve as a satisfactory explanation for hydrogen overvoltage. The tendency is rather to seek

[1] Newbery, J. Chem. Soc., **109**, 1359 (1916).

an explanation in the slowness of some of the steps involved in the formation of hydrogen gas from hydrogen ions in solution.

Before molecular gaseous hydrogen can be formed at any electrode from H_3O^+ ions in solution and escape, the following processes must take place: (1) H_3O^+ ions must diffuse from the solution to the electrode; (2) the H_3O^+ ions must be transferred from the solution to the electrode; (3) the ions on the electrode must be discharged by taking on an electron to yield atomic hydrogen; (4) the atoms of hydrogen must combine to form hydrogen molecules; and (5) the molecules of hydrogen must escape from the electrode as bubbles of gas. If all of these processes could take place with sufficient speed at all electrodes, the latter could in each instance behave reversibly. On the other hand, if one or more of these took place at a rate insufficient to preserve equilibrium at all times, irreversibility would result and through it an overvoltage.

Of the five stages mentioned, the first cannot possibly explain the dependence of overvoltage on the nature of the electrode. If diffusion of H_3O^+ ions to the electrode were the rate-determining process for the overall electrode reaction, the overvoltage would be the same on all metals in a given solution and at a given current density. As this is not the case, this possibility may be dismissed. Again, for various reasons it is highly questionable whether step 5, namely, the escape of molecular hydrogen from the electrode, can be the slowest stage in the electrode reaction sequence.

The slowness of step 4, i.e., the combination of hydrogen atoms to form hydrogen molecules, as the cause of overvoltage was proposed by Tafel.[1] In line with this view is the suggestive fact that the metals on which hydrogen overvoltage is relatively low, such as platinum and palladium, are good catalysts for the combination of hydrogen atoms, whereas metals of high hydrogen overvoltage, such as mercury, lead, tin, cadmium, and bismuth, are poor catalysts for this reaction. Again, on the basis of this theory Tafel was able to show that the overvoltage at constant temperature should vary with the current density according to an equation identical in form with equation (8) actually observed. However, whereas experiment leads to a value of b equal to $2(2.303\ RT)/\mathcal{J}$, Tafel's theory predicts for the slope a value only one-fourth as large. In this respect the theory is not in accord with experiment.

A theory ascribing hydrogen overvoltage to the slowness of discharge of hydrogen ions at the electrode was developed by Smits,[2] Erdey-Gruz and Volmer,[3] Frumkin,[4] and on quantum-mechanical grounds by Gurney[5] and

[1] Tafel, Z. physik. Chem., **50**, 641 (1905).   [2] Smits, *ibid.*, **172A**, 470 (1935).

[3] Erdey-Gruz and Volmer, *ibid.*, **150A**, 203 (1930); Erdey-Gruz and Wick, *ibid.*, **162A**, 53 (1932).

[4] Frumkin, *ibid.*, **164A**, 121 (1933); Acta Physicochim. U.R.S.S., **7**, 475 (1937); *ibid.*, **11**, 21 (1939); *ibid.*, **12**, 481 (1940).   [5] Gurney, Proc. Roy. Soc., **134A**, 137 (1931).

Fowler.[1] In the theory as developed by Frumkin the correct expression is deduced for the slope of the overvoltage-$\log_{10} I$ curve. However, it has been pointed out that the factor 2 in the expression $b = 2(2.303\ RT)/\mathcal{F}$ is arrived at purely by assumption. Again, it has been claimed that the argument employed in this theory may be equally applicable to the transfer of hydrogen ions to the electrode as to the discharge of hydrogen ions as the slow process, and hence the theory is not specific. The quantum-mechanical treatment of Gurney and Fowler leads to the correct result and is not subject to the objections raised against the Frumkin theory. Nevertheless, this approach has been questioned both because it is difficult to see how the combination of an ion and an electron may be slow, and because this theory does not attempt to relate the overvoltage with the nature of the cathode material.

Recently Eyring, Glasstone, and Laidler,[2] using the theory of absolute reaction rates,[3] have proposed a theory of overvoltage in which the slow process leading to hydrogen overvoltage is definitely ascribed to the transfer of protons from the solution to the electrode. Another interesting feature in this theory is the fact that the proton is believed to come not from the hydrogen ions in solution but from *water molecules*. These authors consider the cathode to be covered with a layer of adsorbed water which in turn is surrounded by water molecules in solution. Under the influence of the applied potential protons, i.e., nonsolvated hydrogen ions, are supposed to be transferred from the water molecules in solution to those on the electrode surface. As a result of this transfer a hydroxyl ion is left behind in the solution, while the H_3O^+ ion formed on the surface is rapidly discharged by taking on an electron to from a hydrogen atom. The mechanism postulated can be represented briefly by the scheme

$$HO\text{—}H \quad H_2O \mid \overline{M} \longrightarrow OH^- + H_3O^+ \mid \overline{M} \qquad (10a)$$
$$\text{(Solution)} \qquad \text{(Electrode)} \qquad \text{(Solution)} \qquad \text{(Electrode)}$$

$$H_3O^+ \mid \overline{M} \longrightarrow \frac{H_2O}{H} \bigg| M \qquad (10b)$$
$$\text{(Electrode)} \qquad \qquad \text{(Electrode)}$$
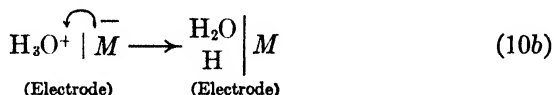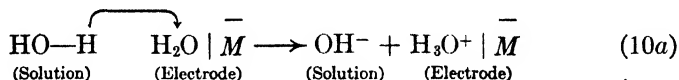
where M is the electrode, and in which the first stage is slow, the second rapid. The slowness of the proton transfer process arises from the presence of a free energy barrier which must be surmounted before a proton can be transferred from the solution to the electrode. This free energy barrier also leads to a potential which becomes included in the

[1] Fowler, Trans. Faraday Soc., **28**, 368 (1932).

[2] Eyring, Glasstone, and Laidler, J. Chem. Phys., **7**, 1053 (1939); Trans. Am. Electrochem. Soc., **76**, 145 (1939).

[3] See next chapter.

overall electrode emf as an overvoltage. Following through the mathematical implications of this theory, Eyring, Glasstone, and Laidler were able to show that, if the overvoltage is independent of pH, it must be related to the current density by equation (9), as observed. Moreover, since the nature of the free energy barrier will depend in general on the ability of the electrode materials to bind protons to themselves, the overvoltage observed should vary with the nature of the electrode, being greatest for electrodes with little affinity for hydrogen and least for those with high affinity. This in general is the case.

Although it is as yet impossible to say to what extent this last theory can encompass the various phenomena observed in connection with hydrogen overvoltage, it does yield the correct dependence of overvoltage on current density, and a connection between overvoltage and electrode material. Again, this theory establishes certain common features between hydrogen and oxygen overvoltage which tend to bring out similarities between the two and to bring them within the compass of a common picture. For these reasons the theory has a good deal to recommend it.

Overvoltage in Anodic Processes. The limited data on anodic deposition indicate that evolution of the halides at electrodes such as platinum and graphite takes place with relatively small overvoltages. Thus, at 100 milliamperes per square centimeter the overvoltages of chlorine, bromine, and iodine are respectively 0.03, 0.01, 0.03 volt at 25° C on platinized platinum, 0.05, 0.01, 0.03 volt on smooth platinum, and 0.25, 0.27, 0.10 volt on graphite. At 1000 milliamperes per square centimeter these rise to about 0.07, 0.20, 0.20 volt on platinized platinum, 0.24, 0.4, 0.22 volt on smooth platinum, and 0.5, 0.33, 0.7 volt on graphite. At very small current densities the overvoltages are practically negligible and indicate, therefore, that the deposition of the halogens takes place essentially reversibly.

However, this is not true of oxygen evolution. Although oxygen overvoltages are difficult to obtain with exactness because of attack of the electrodes by the oxygen, the data available indicate that they are invariably high even at low current densities. This may be seen from Table 2, where the oxygen overvoltages on various metals at several current densities are given. Another important fact revealed by this table is that oxygen overvoltages attain higher values on metals such as platinum and gold than on metals such as copper and nickel, which is just the opposite to the hydrogen overvoltages on these substances. Otherwise oxygen overvoltage behaves essentially analogously to that of hydrogen. In acid solutions the overvoltage appears to be independent of pH, increase of temperature leads to a decrease in overvoltage, and the dependence of overvoltage on current density can again be repre-

sented by equation (9) with exactly the same expression for the slope of the curve. In alkaline solutions, however, the overvoltage appears to vary with the hydroxyl ion concentration.

TABLE 2

OXYGEN OVERVOLTAGES AT 25° C

(Volts)

Current Density (ma/sq cm)	Overvoltage						
	Plat'd Pt	Smooth Pt	Au	Graphite	Cu	Ag	Ni
1	0.40	0.72	0.67	0.53	0.42	0.58	0.35
10	0.52	0.85	0.96	0.90	0.58	0.73	0.52
50	0.61	1.16	1.06	—	0.64	0.91	0.67
100	0.64	1.28	1.24	1.09	0.66	0.98	0.73
500	0.71	1.43	1.53	1.19	0.74	1.08	0.82
1000	0.77	1.49	1.63	1.24	0.79	1.13	0.85
1500	0.79	1.38	1.68	1.28	0.84	1.14	0.87

Theories of Oxygen Overvoltage. There is no question that evolution of oxygen at most anodes is attended with attack of the electrode and attendant coverage of the electrode with a layer of higher oxides of the metal. For a while it was believed that these oxides cause a high back emf, and are thus responsible for oxygen overvoltage. Another theory, in turn, ascribed overvoltage to the accumulation of atomic oxygen at the electrode, while still a third associated oxygen overvoltage with the ability of the anode to catalyze the combination of oxygen atoms to oxygen molecules. In these respects these theories are essentially similar to those proposed to explain hydrogen overvoltage, except that here oxygen is involved rather than hydrogen.

Unlike these theories, that of Eyring, Glasstone, and Laidler claims that the process responsible for oxygen overvoltage is the same as that for hydrogen, namely, the slowness of proton transfer across a free energy barrier existing about the electrode. The only difference is that this time the transfer of the proton is not from a water molecule in solution to a water molecule on the electrode, but from a water molecule on the electrode to one in solution. In other words, the process now is the slow step,

$$
\overbrace{\text{H}_2\text{O} \quad \text{H--O}}\,\Big|\, + \quad\quad\quad\quad +
$$
$$
\text{H}_2\text{O} \quad \text{H--O} \Big|\!\!\begin{array}{c} \\ | \\ \text{H} \end{array}\!\!\Big| M \longrightarrow \text{H}_3\text{O}^+ + \text{OH}^- \,|\, M \qquad (11a)
$$

(Solution)   (Electrode)   (Solution)   (Electrode)
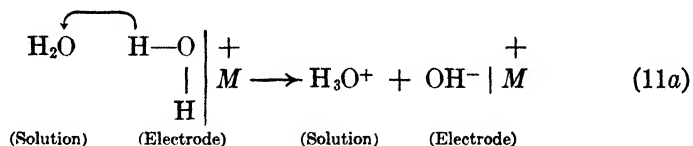
which results in the formation of a hydrogen ion in solution and a hydroxyl ion on the electrode. This slow step is followed by the rapid discharge and combination of hydroxyl groups on the electrode to yield

$$\overset{+}{OH^-} \mid M \longrightarrow OH \mid M \qquad (11b)$$
(Electrode)   (Electrode)

$$2\ OH \longrightarrow H_2O + O \qquad (11c)$$
$$2\ O \longrightarrow O_2(g) \qquad (11d)$$
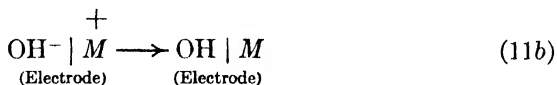
From an analysis of this mechanism in terms of the theory of absolute reaction rates equation (9) again follows for the variation of oxygen overvoltage with current density. Again, since electrodes which exhibit low affinity for hydrogen will exhibit high affinity for hydroxyl ions, the oxygen overvoltage barrier should be lower on metals with high hydrogen overvoltage than those with low. As a consequence it may be anticipated that metals with high hydrogen overvoltage will have low oxygen overvoltage, and vice versa. This expectation is borne out to a degree by the experimental data, although exact parallelism is difficult to establish because of oxide formation and other complications observed in oxygen evolution at various anodes.

Metal Deposition and Hydrogen Overvoltage. Reference has already been made to the fact that deposition of most metals from aqueous solutions takes place with very little overvoltage, and hence at essentially reversible potentials. Were the metal ions the only cations present, these potentials would alone determine the deposition. However, in any aqueous solution of the metal hydrogen ions are also present, and consequently two cathodic reactions are possible, the deposition of the metal and the evolution of hydrogen. The question arises, therefore: Under what conditions can a metal be deposited from any aqueous solution before evolution of hydrogen sets in?

The answer to this question is simply this: If the potential required for metal deposition is less than that for the hydrogen, the product of electrolysis at the cathode will be the metal. If the reverse is true, cathodic liberation of hydrogen will occur. The whole problem reduces itself, therefore, to a matter of the potentials of the respective processes. In this connection it should be remembered that, whereas the cathodic deposition of metals involves essentially only the reversible metal electrode potentials, the liberation of hydrogen involves not only the reversible potential of the hydrogen electrode in the particular solution, but also the overvoltage of hydrogen at the material composing the cathode. In fact, it is largely because of the latter factor that many metals can be deposited from aqueous solutions under conditions which in absence of overvoltage would invariably yield hydrogen.

To illustrate the above points, consider first the electrolysis of an aqueous solution of silver nitrate in which the activity of the salt is unity. In this solution the deposition potential of the silver will be the standard electrode potential of the metal, or $\mathcal{E}^0_{Ag} = + 0.80$ volt at 25° C. Again, in this same solution the activity of hydrogen ions is 10^{-7}, and hence the potential for the reversible evolution of hydrogen would be

$$\mathcal{E}_{H_2} = - \frac{.059}{1} \log_{10} \frac{1}{10^{-7}}$$
$$= - 0.41 \text{ volt}$$

Since the deposition of silver requires a lesser expenditure of energy than the evolution of hydrogen, as the former process is spontaneous whereas the latter is not, silver will necessarily plate out in preference to hydrogen. In this case, therefore, silver will be the electrolytic product at the cathode whether hydrogen overvoltage is present or not. The same will be true also of deposition of copper. Further, although the difference between the silver or copper potentials and that of hydrogen can be reduced by making the solutions acid, still the potential of the latter cannot be reduced below those of the metals mentioned, and these will plate out whether the solution is acid or neutral.

On the other hand, the deposition of most metals above hydrogen in the electromotive series would be impossible were it not for overvoltage. For a solution of cadmium ions at unit activity the potential of the cadmium electrode will be 0.40 volt, and hence the potential required for deposition will be $- 0.40$ volt. This potential is so close to that of the hydrogen electrode in neutral solution, $- 0.41$ volt, that both cadmium and hydrogen may be expected to deposit. Again, if the solution were acid, the potential of the hydrogen electrode would be less negative, and in such solutions it would appear to be impossible to plate cadmium at all. Still, this is not the case, for evolution of hydrogen does not take place reversibly on cadmium cathode; in fact, at a current density as low as 1 milliampere per square centimeter the overvoltage of hydrogen on cadmium is already about a volt. Consequently, in order to plate hydrogen from neutral solutions of a cadmium salt a potential of about $- 1.4$ volts will be necessary, while from strongly acid solutions about $- 1.0$ volts. As both of these potentials are considerably higher than that required for cadmium plating, $- 0.40$ volt, the latter process takes place rather than hydrogen evolution.

In this manner may be explained the electrolytic separation of such metals as zinc, iron, nickel, tin, and lead from their aqueous solutions. Were it not for the presence of hydrogen overvoltage, the deposition of some of these metals would be complicated by appearance of hydrogen, while in other cases, as with zinc, hydrogen would be the exclusive

product. Particularly illustrative in this connection is the fact that even sodium may be deposited on a mercury cathode in preference to hydrogen. With an overvoltage of ca. a volt on mercury, a potential of about − 1.4 volts is required to liberate hydrogen on this metal. On the other hand, deposition of sodium requires a potential of over 2 volts. However, when sodium is deposited on mercury, the former dissolves in the latter to form an amalgam in which the activity of the sodium is moderated. When the amalgam is dilute, the potential of the sodium in it exerts a potential of only about 1.2 volts against a solution of sodium ions instead of over 2 volts for pure sodium, and hence the deposition of sodium on the amalgam can take place at a potential of − 1.2 volts. As this potential is less than the − 1.4 volts for hydrogen on mercury, sodium can be recovered from aqueous solutions as an amalgam in preference to hydrogen. This fact is utilized commercially in the Castner cell for the electrolysis of sodium chloride, where the sodium is recovered first cathodically as an amalgam, and is allowed then to react with water to yield hydrogen and sodium hydroxide.

Electrolytic Separation of Metals. Not only may metals be separated electrolytically from hydrogen ions in solution when sufficient differences in potential exist, but also various metals from each other. To accomplish clean-cut separation of one metal from another it is desirable that a difference of about 0.2 volt exist between the potential of one metal at near complete removal and the potential of the next metal at the inception of its separation. When such conditions obtain, the various metals can be removed from solution and separated from each other quantitatively with sufficient accuracy for analytical purposes. In this manner silver may be separated from copper, copper from cadmium, and cadmium from zinc. In a solution containing all of the metals mentioned the procedure is to adjust the electrolytic potential first at a point higher than that required to deposit silver, but not high enough to yield copper. At such a potential appreciable current will flow as long as any silver is present in solution, and deposition takes place. However, as soon as the silver is removed the potential is not sufficiently high to permit the next metal in the series to plate, and hence the current falls to a low value, indicating the end of the silver separation. At this stage the cathode is removed, a new one is inserted, the potential is adjusted to a value intermediate between the copper and cadmium, and electrolysis is again initiated. By repeating this sequence of potential adjustment and electrolysis, the copper can be removed next, then the cadmium, and finally the zinc.

Occasionally, in order to avoid gassing at the cathode due to hydrogen evolution, the acidity of the solution may require adjustment before separation of the metals higher in series. Again, frequently a greater

difference in separation potentials can be obtained by converting some of the ions to complexes, or by changing the temperature. Various such artifices and other experimental details may be found described in books on quantitative electro-analysis.

Simultaneous Deposition of Metals. In order to deposit two metals simultaneously at a cathode it is necessary that the two have the same deposition potentials. This means that the concentrations of the two ions in solution must be so adjusted that the potentials of the two metals, which are ordinarily different, are brought together. Such adjustment is usually not possible with solutions of simple salts of the metals, but may be accomplished by converting the metal ions to complexes. The outstanding example of this type of behavior is the direct formation of brasses by simultaneous deposition of copper and zinc. In aqueous solutions of their simple salts these two metals have deposition potentials that differ by over a volt. However, when the cations of these metals are converted to their cyanide complexes, the difference in potential between them is considerably reduced until at a given concentration of cyanide they become essentially the same. From such solutions, therefore, both metals can be plated simultaneously as a brass. Furthermore, by varying the proportions of zinc, copper, and cyanide in the electrolyte, and by controlling the current density, deposits can be obtained varying in metal content from pure copper to pure zinc; i.e., brasses of any desired composition can be obtained. It is also of interest to point out that various addition agents and the temperature alter the composition of the deposit. Increase in temperature invariably leads to an increase in copper content, so that at sufficiently high temperatures the product is exclusively this metal.

The Anodic Solution of Metals. When an inert anode, such as platinum, is used in the electrolysis of a solution which can yield only oxygen, the product at the anode is always this gas. On the other hand, when a more active metal is the anode material, reaction at the anode may be either solution of the metal or evolution of oxygen, depending on which of these processes requires the lower potential. However, since the oxygen overvoltage is high on most metals, only platinum and gold will usually yield oxygen, while other metals, such as copper, silver, cadmium, tin, and zinc, will dissolve anodically to yield ions of the metals.

During the anodic solution most metals exhibit their equilibrium potentials in the given solution, and hence the potential required for the process need be only slightly more positive than that of the metal itself. With iron, nickel, and cobalt, however, potentials exceeding the equilibrium value by 0.3 or 0.4 volt are required before solution commences, indicating that the solution of these metals is not reversible, but is

attended by appreciable overvoltage. Nevertheless, as long as the requisite potential is applied and the current density is low, the metals continue to pass into solution. But if the current density is raised, a point is reached at which the potential of the metal rises rapidly, the current through the solution falls, and the metal stops dissolving. Any attempt to increase the current by raising the voltage does not rectify the situation. Finally at a sufficiently high potential oxygen evolution sets in, and only then does the current begin to go up again.

This suspension of activity and solubility of the metals is called *passivity*, and in this state the metals are said to be *passive*. Of the three metals mentioned, nickel is rendered passive more readily than cobalt, and cobalt more readily than iron under identical conditions. Again, these metals become passive more readily in alkaline than in acid solutions. In these metals passivity is promoted by presence of oxidizing agents in the electrolyte, while halide ions tend to hinder it. On the other hand, molybdenum and tungsten are passivated more easily in acid than in alkaline solutions. In general it may be said that a metal can be passivated more readily in a solution in which its ions are precipitated the more easily. Thus, iron, cobalt, and nickel cations are soluble in acid solutions but precipitate in alkaline solutions as the hydroxides, whereas molybdenum and tungsten are soluble in alkalies, but precipitate readily in acid solutions.

A metal once rendered passive will regain its activity slowly on standing after the current has been shut off. Reactivation can be accelerated appreciably by making the metal a cathode in an electrolytic cell, or by contacting it with a metal more active than itself. Scratching the surface also promotes the return of activity.

The passivity of metals has been traced to the formation of oxide films on the surfaces of the metals. These films are extremely thin, cannot ordinarily be seen, and are themselves not soluble. By covering the surface of the metal, the film of oxide can prevent the solution of the more active metal underneath, and can thus make the metal behave as if it were much more noble than it actually is.

Electrolytic Oxidation and Reduction. Deposition of metallic ions or hydrogen on the cathode, and deposition of oxygen or solution of metals at the anode, are not the only processes which may take place at these electrodes. The fact is that any reaction leading to a removal of electrons from the cathode, resulting thus in a reduction, should be possible at the cathode, whereas any reaction leading to a giving up of electrons, and hence an oxidation, should be possible at the anode. This is actually the case. When such reductions and oxidations can take place at potentials lower than those required for evolution of oxygen or hydrogen, formation and evolution of these gases is avoided, and instead

reduction of the substances present in solution takes place *at the cathode*, *oxidation at the anode.* Substances which when present in solution can thus undergo oxidation or reduction, and at the same time prevent evolution of hydrogen or oxygen, are called *depolarizers.* Further, those substances which by their own reduction can prevent formation of hydrogen are called *cathodic depolarizers*, while those which by their own oxidation can prevent evolution of oxygen are referred to as *anodic depolarizers.*

The simplest depolarizers are inorganic ions capable of existence in solution in two valence states. Thus ferric and stannic ions by reduction at the cathode to ferrous and stannous ions can depolarize the evolution of hydrogen, while the oxidation of ferrous or stannous ions at the anode can depolarize the formation of oxygen. In these instances there is evidently no deposition of gas, and the electrode processes involve apparently direct electron transfer between ions and electrode. Furthermore, such processes involve the establishment of a definite potential at the electrode dependent on the ratio of the activities of the ions, although occasionally overvoltage is encountered.

In distinction to these simple depolarizers are the inorganic and organic depolarizers which involve apparently the evolution of hydrogen or oxygen and subsequent reaction with these gases. Thus the overall processes in the reduction of nitrates and nitrobenzene at the cathode are

$$NaNO_3 + 2\,H = NaNO_2 + H_2O$$
$$C_6H_5NO_2 + 6\,H = C_6H_5NH_2 + 2\,H_2O$$
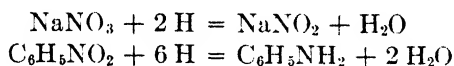
while the anodic oxidation of alcohol to acetic acid is given by

$$C_2H_5OH + 2\,O = CH_3COOH + H_2O$$
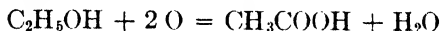
Such processes as a rule do not involve a definite potential for the system, and are usually irreversible. Their efficacy arises from the fact that by reaction with any oxygen or hydrogen deposited on an electrode they can prevent accumulation and escape of these gases, and operate thus to permit the electrode processes at lower potentials than would be required to form the gases.

Besides the electrolytic reductions mentioned, many others can be carried out and are utilized industrially. Among these may be quoted the reduction of permanganate to manganate, chromate to chromic salts, quinone to hydroquinone, sugars to alcohols, and of nitrobenzene to aniline and various other intermediate products. In the latter reduction high hydrogen overvoltage cathodes, such as lead, zinc, copper, tin, and mercury, invariably yield the product of most complete reduction, aniline. On the other hand, on nickel, carbon, silver, or platinum

cathodes in neutral solution the product is mostly C_6H_5NHOH. Again, in acid solutions on these cathodes p-aminophenol $\left(C_6H_4 \underset{OH}{\overset{NH_2}{\diagdown}} \right)$, benzidine ($NH_2C_6H_4$—$C_6H_4NO_2$), as well as aniline can be recovered, whereas in alkaline solutions azoxybenzene $\left(\begin{array}{c} C_6H_5N—NC_6H_5 \\ \diagdown\diagup \\ O \end{array} \right)$ is the final reduction product. By varying the conditions of electrolysis, therefore, a multiplicity of reduction products can be obtained from a single depolarizer, nitrobenzene.

Many and varied types of anodic oxidations have been carried out electrolytically, with quite a number of these of definite industrial importance and application. Among these may be mentioned the oxidation of ferrous to ferric salts, stannous to stannic salts, of manganate to permanganate, of chromium salts to chromates, and of ferro- to ferricyanides. Interesting types of oxidations are the conversion of carbonates to percarbonates,

$$2\ CO_3^{--} \longrightarrow C_2O_6^{--}$$

of sulfates to persulfates,

$$2\ SO_4^{--} \longrightarrow S_2O_8^{--}$$

and of borax ($Na_2B_4O_7$) to sodium perborate ($NaBO_3 \cdot 4\ H_2O$). The persulfates on heating with water yield hydrogen peroxide, and this is the electrolytic method for preparing this substance. Finally, electrolytic oxidations of organic compounds are exemplified by the oxidation of hydroquinone to quinone, of alcohol to acetic acid, and of salts of saturated chain organic acids to hydrocarbons and carbon dioxide (Kolbe's reaction)

$$2\ RCOO^- = \underset{\underset{RH}{|}}{RH} + 2\ CO_2$$
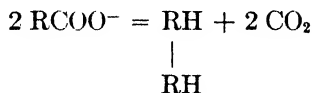
By this means hydrocarbons containing up to 34 carbon atoms have been prepared.

Commercial Cells. We shall conclude this chapter with a description of several electrochemical cells commonly employed as sources of electrical energy. Strictly speaking, the discussion of these cells belongs in the preceding chapter; however, certain features about them can be understood more readily at this junction.

The cells employed generally as sources of electric potential are of two types. In the *primary cells* electrical energy can be obtained at the

expense of chemical reactivity only as long as the active materials are still present. Once these have been consumed, the cell cannot be profitably and readily rejuvenated or recharged, and must be discarded. On the other hand, in *secondary* or *storage cells* the cell once used can be recharged by passage of current through it, and it may hence be used over and over again. Of cells of the first type, the only one of importance is the Leclanché or dry cell, while among the storage type the acid lead cell and the alkaline Edison cell are of common occurrence. These cells will now be described in turn.

(a) *The Leclanché or Dry Cell.* The Leclanché primary cell consists of a negative zinc electrode, a positive carbon electrode surrounded by manganese dioxide, and a paste of ammonium chloride and zinc chloride as an electrolyte. The reaction at the negative electrode is solution of zinc to form zinc ions, while that at the positive electrode appears to be

$$2\,MnO_2 + H_2O = Mn_2O_3 + 2\,OH^- \tag{12}$$
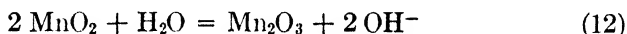
The exact mechanism at the latter electrode is not clear, and various schemes have been proposed to account for its action. The simplest of these is possibly the sequence,

$$2\,H_2O = 2\,H^+ + 2\,OH^-$$
$$2\,H^+ + 2\,\ominus = 2\,H$$
$$2\,MnO_2 + 2\,H = Mn_2O_3 + H_2O$$
$$2\,MnO_2 + H_2O = Mn_2O_3 + 2\,OH^- \tag{13}$$
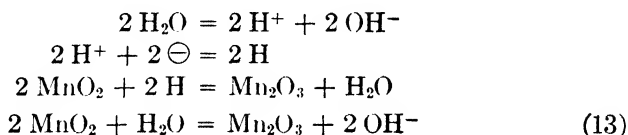
in which the function of manganese dioxide is to act as a depolarizer for the deposition of hydrogen on the graphite electrode. From these two electrode reactions the overall process for the cell follows as

$$Zn + 2\,MnO_2 + H_2O = Zn^{++} + 2\,OH^- + Mn_2O_3 \tag{14}$$
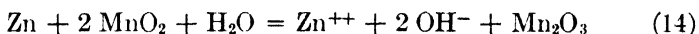
The hydroxyl ions generated by the action of the cell liberate ammonia from the ammonium chloride, which in turn combines with the zinc ions to precipitate the difficultly soluble salt $Zn(NH_3)_2Cl_2$

$$2\,NH_4Cl + 2\,OH^- = 2\,NH_3 + 2\,Cl^- + 2\,H_2O \tag{15a}$$
$$Zn^{++} + 2\,NH_3 + 2\,Cl^- = Zn(NH_3)_2Cl_2 \tag{15b}$$
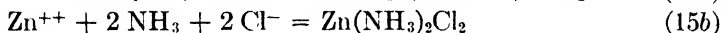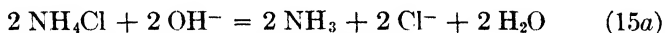
However, reactions (15a) and (15b) are secondary processes not involved directly in the electrode reactions, so they do not contribute to the potential of the cell.

The Leclanché cell develops a potential of about 1.5 volts on open circuit. With usage this voltage tends to fall until eventually the cell has to be discarded. At best these cells have a low capacity. They are not intended for heavy duty.

(b) *The Lead Storage Cell.* In this cell the negative electrode is lead, the positive electrode lead impregnated with lead dioxide, while the electrolyte is a solution of approximately 20 per cent sulfuric acid with a specific gravity of about 1.15 at room temperature. On drawing current from this cell lead dissolves at the negative electrode to form lead ions, which combine then with sulfate ions in solution to precipitate lead sulfate, namely,

$$Pb(s) = Pb^{++} + 2 \ominus \qquad (16a)$$
$$Pb^{++} + SO_4^{--} = PbSO_4(s) \qquad (16b)$$
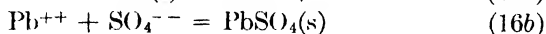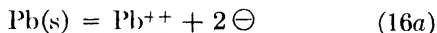
The overall electrode process for 2 faradays is, therefore,

$$Pb(s) + SO_4^{--} = PbSO_4(s) + 2 \ominus \qquad (16c)$$

which shows that actually this electrode is a lead-lead sulfate electrode reversible to sulfate ions. The action of the positive electrode, in turn, can be accounted for by postulating that the lead dioxide by hydration can yield plumbic ions,

$$PbO_2(s) + 2 H_2O = Pb^{++++} + 4 OH^- \qquad (17a)$$
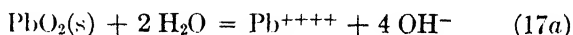
which are then reduced to plumbous ions by taking on two electrons from the electrode:

$$Pb^{++++} + 2 \ominus = Pb^{++} \qquad (17b)$$
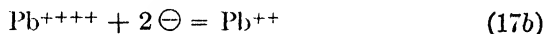
Since the plumbous ions are in solution, they react with sulfate ions to precipitate lead sulfate,

$$Pb^{++} + SO_4^{--} = PbSO_4(s) \qquad (17c)$$

while the hydroxyl ions formed in reaction (17a) are neutralized by the acid:

$$4 OH^- + 4 H^+ = 4 H_2O \qquad (17d)$$
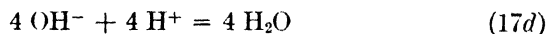
On adding the last four reactions, the overall process at the lead dioxide electrode becomes

$$PbO_2(s) + 4 H^+ + SO_4^{--} + 2 \ominus = PbSO_4(s) + 2 H_2O \qquad (17e)$$
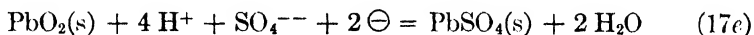
from which we see that this is essentially an oxidation-reduction electrode dependent on both hydrogen and sulfate ions for its potential. From equations (16c) and (17e) the overall reaction in the lead storage cell on discharge follows as

$$Pb(s) + PbO_2(s) + 2 H_2SO_4 = 2 PbSO_4(s) + 2 H_2O \qquad (18)$$
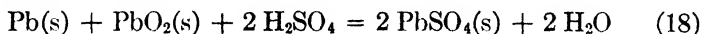
According to equation (18) the potential of the lead storage cell should depend at any given temperature only on the activity of sulfuric acid in solution. This is actually the case. Thus, at 25° C and at a concentration of 7.4 per cent sulfuric acid the potential is 1.90 volts, at 21.4 per cent 2.00 volts, and at 39.2 per cent 2.14 volts. Further, since sulfuric acid is consumed in the operation of the cell, the specific gravity of the

electrolyte should fall on discharge, as it does. Finally, the heats of reaction (18) for different concentrations of sulfuric acid obtained electrochemically from the Gibbs-Helmholtz equation are in excellent accord with thermally obtained results, confirming thus, if not the mechanism of the electrode processes, at least the validity of the overall cell reaction (18).

To recharge a lead storage cell, reaction (18) taking place on discharge must be reversed. For this purpose a potential higher than that of the cell must be applied externally. During this electrolysis lead is deposited on the cathode, lead dioxide is formed at the anode, and sulfuric acid is regenerated in the cell. Owing to this increase in acid concentration the potential of the cell rises, and so does the specific gravity. As long as the lead ions are present in solution, no gassing will occur on charging. However, if the electrolysis is permitted to proceed until all the lead is removed, i.e., the battery is overcharged, evolution of hydrogen at the lead surface and oxygen at the lead dioxide surface will take place, and there is danger of injury to the electrode plates. For this reason lead cells should be charged only until the specific gravity is up to the desired point and no further.

The voltage of a lead cell does not depend on the size of the cell or of the electrodes. Small or large, a lead cell always has the same potential for a given acid concentration. The only advantage of large electrodes over small is that they give a cell a larger capacity to deliver electrical energy. It may also be pointed out that a lead cell is fairly efficient, enabling recovery on discharge of 90 to 95 per cent of the quantity of electricity put into the cell on charge.

(c) *The Edison Cell.* The Edison storage cell consists of a negative electrode of iron and a positive electrode of nickel in contact with the sesquioxide of nickel (Ni_2O_3). The electrolyte is a 21 per cent by weight solution of potassium hydroxide to which is added some lithium hydroxide. Just exactly what is the function of the latter is not well understood, but it seems to improve the capacity of the cell. On discharge, the electrode reactions at the iron surface can be represented by

$$Fe(s) = Fe^{++} + 2 \ominus \qquad (19a)$$
$$\underline{Fe^{++} + 2\,OH^- = FeO(s) + H_2O} \qquad (19b)$$
$$Fe(s) + 2\,OH^- = FeO(s) + H_2O + 2 \ominus \qquad (19c)$$
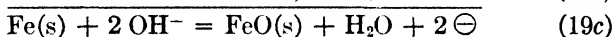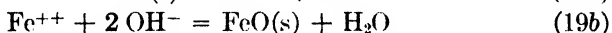
those at the positive electrode by

$$Ni_2O_3(s) + 3\,H_2O = 2\,Ni^{+++} + 6\,OH^- \qquad (20a)$$
$$2\,Ni^{+++} + 2 \ominus = 2\,Ni^{++} \qquad (20b)$$
$$\underline{2\,Ni^{++} + 4\,OH^- = 2\,NiO(s) + 2\,H_2O} \qquad (20c)$$
$$Ni_2O_3(s) + H_2O + 2 \ominus = 2\,NiO(s) + 2\,OH^- \qquad (20d)$$
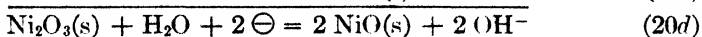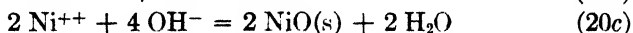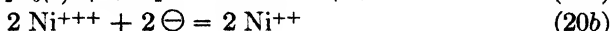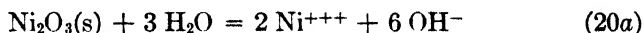
From these the overall cell reaction on discharge follows as

$$Fe(s) + Ni_2O_3(s) = FeO(s) + 2 NiO(s) \qquad (21)$$
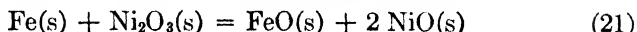
while on charge it is the reverse of this. Equation (21) indicates that the voltage of the cell should be independent of the concentration of the electrolyte. This, however, is not quite the case, for the potential does depend slightly on the concentration of the base. This variation may be accounted for by postulating that the various oxides involved in equation (21) are hydrated, in which instance water would appear in the cell reaction. As the activity of the water changes with the concentration of potassium hydroxide, so would the potential of the cell.

The potential of the Edison cell is lower than that of the lead cell, being only about 1.3 volts. Its efficiency is also lower. On the other hand, its lower weight compared to the lead cell, and the absence of acid, recommend it for certain purposes.

REFERENCES FOR FURTHER READING

1. C. J. Brockman, "Electrochemistry — Principles and Practice," D. Van Nostrand Company, Inc., New York, 1931.
2. C. J. Brockman, "Electro-Organic Chemistry," John Wiley & Sons, Inc., New York, 1926.
3. Creighton and Koehler, "Principles and Applications of Electrochemistry," 2 volumes, John Wiley & Sons, Inc., New York, 1935.
4. S. Glasstone, "Introduction to Electrochemistry," D. Van Nostrand Company, Inc., New York, 1942.
5. Glasstone and Hickling, "Electrolytic Oxidation and Reduction," D. Van Nostrand Company, Inc., New York, 1935.
6. C. L. Mantell, "Industrial Electrochemistry," McGraw-Hill Book Company, Inc., New York, 1940.
7. M. deK. Thompson, "Theoretical and Applied Electrochemistry," third edition, The Macmillan Company, New York, 1939.
8. G. W. Vinal, "Storage Batteries," third edition, John Wiley & Sons, Inc., New York, 1940.

PROBLEMS

1. Which of the following cells, as *written*, are electrolytic and which are primary cells:
    (a) $Cd(s) \mid CdSO_4(a = 1), PbSO_4(s) \mid Pb(s)$
    (b) $Au(s) \mid H_2SO_4$ (dilute solution) $\mid Pt(s)$
    (c) $Pt(s) \mid Na_2SO_4$ solution $\mid Pt(s)$
    (d) $Ag(s) \mid AgNO_3(m = 0.5) \parallel AgNO_3(m = 0.005) \mid Ag(s)$
    (e) $Pb(s) \mid Pb^{++}(a = 1) \parallel Ag^+(a = 1) \mid Ag(s)$

2. Write the electrode and cell reactions which will occur in the above cells when a current is passed through them.

3. A 2.00 volt storage battery is used in an electrolysis in which a back emf of 1.45 volts is developed. If the resistance of the entire circuit is 10 ohms, (a) what is the magnitude of the current, and (b) what quantity of heat is generated by the current flow per faraday of electricity?

Ans. (a) 0.055 amp; (b) 12,690 cal

4. Under reversible conditions what applied potential will be necessary to initiate steady electrolysis in each of the following cells:

    (a) $Pb(s) \mid PbSO_4(s), H_2SO_4, PbSO_4(s) \mid Pb(s)$

    (b) $Pt(s) \mid Fe^{++}(a = 1), Fe^{+++}(a = 1) \parallel Br^-(a = 1), AgBr(s) \mid Ag(s)$

    (c) $Ni(s) \mid NiSO_4(a = 1), PbSO_4(s) \mid Pb(Hg)$

    (d) $Ag(s) \mid AgCl(s), ZnCl_2(a = 1) \mid Zn(s)$

The $\mathcal{E}^0$ value for the $Pb(Hg) \mid PbSO_4(s)$ $SO_4^{--}(a = 1)$ electrode is 0.3505 volt.

5. Repeat the preceding problem for each of the following cells:

    (a) $Pt(s) \mid HI(a = 1) \mid Pt(s)$

    (b) $Pt(s) \mid CdCl_2(a = 1) \mid Pt(s)$

    (c) $Pt(s) \mid NiSO_4(a = 1) \mid Pt(s)$

6. Calculate the reversible decomposition potentials of the following cells:

    (a) $Pt(s) \mid HBr(m = 0.05, \gamma_\pm = 0.860) \mid Pt(s)$

    (b) $Ag(s) \mid AgNO_3(m = 0.5, \gamma_\pm = 0.526) \parallel AgNO_3(m = 0.01,$

$$\gamma_\pm = 0.902) \mid Ag(s)$$

    Ans. (a) $- 1.2276$ volts; (b) $- 0.0867$ volt

7. In the cell

$$Pt(s) \left| \begin{array}{l} CdCl_2(m\gamma_\pm = 1) \\ NiSO_4(m\gamma_\pm = 1) \end{array} \right| Pt(s)$$
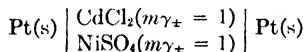
what will be the cell reaction and the reversible decomposition potential when electrolysis is initiated?

8. The overvoltage of H_2 on Pb was determined in a 0.1 molal H_2SO_4 solution ($\gamma_\pm = 0.3$) by measuring during electrolysis the potential of the lead cathode against a normal calomel electrode. The potential thus observed was 1.0650 volts. What is the overvoltage of H_2 on Pb? *Ans.* 0.713 volt

9. The overvoltages of H_2 and O_2 on Ag at a current density of 0.1 amp/square centimeter are, respectively, 0.87 and 0.98 volt. What will be the decomposition potential of a dilute $NaOH$ solution between Ag electrodes at a current density of 0.1 amp/square centimeter?

10. (a) Using the data of Table 2, plot overvoltage against log of the current density for a Ag cathode, and evaluate the constants a and b in equation (8). (b) At what current density will the overvoltage be 0.80 volt?

    Ans. (a) $a = 0.58$, $b = 0.19$; (b) 15.5 mamp/cm²

11. The overvoltage of H_2 on polished Pt is 0.24 volt when the current density is 1 milliampere/square centimeter. What will be the deposition potential of H_2 on this electrode from a solution of pH = 3?

12. The overvoltage of H_2 on Fe is 0.40 volt at a current density of 1 milliampere/square centimeter. In a solution containing Fe^{++} ions at unit activity, what is the lowest pH at which deposition of Fe is possible under these conditions?

    Ans. 0.69

13. The overvoltage of H_2 on Zn is 0.72 volt. If it is desired to deposit Zn^{++} down to a concentration of 10^{-4} moles/liter without deposition of H_2 occurring, what must be the lowest permissible pH of the solution?

14. A solution contains Fe^{++} and Zn^{++} ions at unit activity. If the overvoltage of H_2 on Fe is 0.40 volt, what must be the maximum pH of the solution in order to plate Fe but obtain "gassing" at the cathode prior to Zn deposition? What will be the activity of Fe^{++} ions when "gassing" begins?

    Ans. pH $= 6.10$; $a_{Fe^{++}} = 1.55 \times 10^{-11}$

15. A solution contains Fe^{++} ions at unit activity. (a) Under reversible conditions what must be the activity of Cd^{++} ions in the same solution in order to obtain simultaneous deposition of Fe and Cd? (b) What would have to be the activity of Ni^{++} ions to obtain Ni and Fe simultaneously?

16. A solution contains $Ag^+(a = 0.05)$, $Fe^{++}(a = 0.01)$, $Cd^{++}(a = 0.001)$, Ni^{++} $(a = 0.1)$, and $H^+(a = 0.001)$. The overvoltage of H_2 on Ag is 0.20 volt, on Ni 0.24 volt, on Fe 0.18 volt, on Cd 0.30 volt. Predict what will happen at a cathode immersed in this solution as the potential applied is gradually increased from zero.

17. A solution containing Zn, Na, and H_2 ions at unit activity is electrolyzed with a Hg cathode. Explain what will occur at the cathode as the potential is gradually increased.

18. The overvoltage of O_2 on a Pt anode is 0.72 volt. From a solution of pH = 7 and containing Cl^- of 0.1 activity, which ion will be deposited first?

19. A solution is made originally 0.1 molar in Ag^+ ions and 0.25 molar in KCN. If the dissociation constant of $Ag(CN)_2^-$, $K = [Ag^+][CN^-]^2/[Ag(CN)_2^-]$, is 3.8×10^{-19}, what will be the concentration of Ag ions in this solution, and what will be the deposition potential of Ag? For this calculation assume activities equal to concentrations. *Ans.* 1.52×10^{-17}; $- 0.186$ volt

20. A solution is 0.1 molar in Au^+ and 0.1 molar in Ag^+ ions. The standard reduction potential of Au^+ ($Au^+ + \ominus = Au$) is 1.68 volts, while the dissociation constant of $Au(CN)_2^-$, $K = [Au^+][CN^-]^2/[Au(CN)_2^-]$, is 5×10^{-39}. Find what concentration of NaCN would have to be added to the given solution in order to deposit Au and Ag simultaneously. *Ans.* 0.2 molar

Kinetics of Homogeneous Reactions

Introduction. *Chemical kinetics* is the branch of physical chemistry which concerns itself with the study of the velocity of chemical reactions and with the elucidation of the mechanism by which they proceed.

A question of importance which thermodynamics does not touch upon is: *How rapidly and by what mechanism does a reaction take place?* Thermodynamics considers only the energy relations between reactants and the products of the reaction, but makes no attempt to indicate the stages through which the reactants may have to pass to reach the final products. Above all, it is not concerned with the rate at which equilibrium is attained. In this respect chemical kinetics complements thermodynamics by supplying information about the rate of approach to equilibrium, and also, whenever possible, about the mechanism responsible for the conversion of reactants to products.

Not all reactions lend themselves readily to kinetic study. Many ionic reactions, such as ionizations, neutralizations, and precipitations, are so rapid that to all appearances they are instantaneous. Explosions and certain other reactions, of which $N_2O_4 \longrightarrow 2\,NO_2$ is an example, also proceed so quickly that either it is impossible to determine their rate, or special methods have to be devised to do so. On the other hand, some reactions are so slow that months and even years are necessary to observe any perceptible transformation at ordinary temperatures. Between these extremes lie the reactions whose velocities are measurable. In this category fall many gaseous reactions, as well as many reactions in solution involving both organic and inorganic substances.

The rates of various reactions depend not only on the nature of the reacting substances but also on the temperature and the concentrations of the reactants. An increase in temperature leads to an increase in velocity; in fact, for many reactions an increase of 10° C in temperature may increase the rate of reaction twofold or more. Again, with the exception of certain (zero order) reactions, upon which concentration is without effect, an increase in the initial concentration of reactants also results in an acceleration of the rate. However, the rate for any

given initial concentration does not remain constant during the course of the reaction, but is highest at the beginning, when the concentration is a maximum, and falls with time as the reactants are consumed. In general the variation of the rate of a reaction with time is represented by a curve such as that shown in Fig. 1. This curve indicates that the rate decreases rapidly in the initial stages of the reaction, then much more slowly, and approaches zero asymptotically to the time axis. Theoretically infinite time would be required for the rate to fall to zero. Practically, however, the rate becomes so low after a time that to all intents and purposes the reaction may be considered completed within a finite time interval.
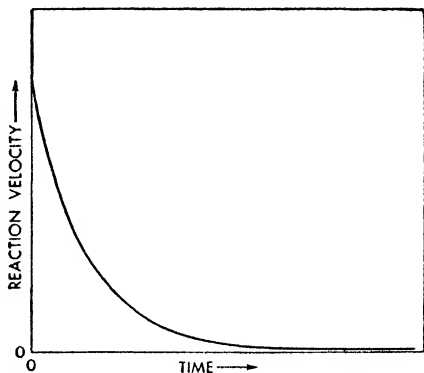
Fig. 1. Variation of Reaction Rate with Time

Further, many reactions are influenced by the presence of substances which possess the ability of accelerating or decelerating the rates of such reactions. Substances possessing this property of affecting the rates of chemical reactions are called *catalysts*, while the reactions thus affected are *catalyzed reactions*. Again certain reactions are uninfluenced by light, while others, called *photochemical reactions*, are greatly stimulated when light of appropriate frequency is permitted to pass through the reacting mixture.

Kinetic studies of reactions are generally carried out at constant temperature. A reacting mixture of known composition is prepared, thermostated, and the decrease in concentration of reacting substances with time is observed by some suitable means. From the time-concentration data thus obtained the kinetic behavior of the reaction can be deduced with the aid of certain principles which will be developed shortly. By repeating this procedure at a number of temperatures, the temperature coefficient of the reaction can be obtained, and from this the dependence of the rate of a reaction on the temperature.

Reactions may be classified kinetically as being either *homogeneous* or *heterogeneous*. A reaction is said to be kinetically *homogeneous* if it takes place in one phase only. If two or more phases are involved in the process, as in a gaseous reaction proceeding on the surface of a solid catalyst or on the walls of the container, the reaction is said to be *heterogeneous*. The principles governing the kinetics of some heterogeneous reactions will be discussed in the next chapter. Here attention will

be devoted to three types of homogeneous reactions, namely: (a) uncatalyzed reactions, (b) catalyzed reactions, and (c) chain reactions. Because of space limitations photochemical reactions will not be discussed in this book.

Order and Molecularity of Reactions. By the rate of a chemical reaction is understood the *rate at which the concentrations of reacting substances vary with time*, i.e., $- dC/dt$, where C is the concentration of reactant while t is the time. The minus sign is used to denote that the rate is negative, in other words, that it decreases with time. According to the law of mass action the rate of any reaction should at each instant be proportional to the concentrations of the reactants, with each concentration raised to a power equal to the number of molecules of each specie participating in the process. Thus, for the reaction

$$A \longrightarrow \text{Products} \qquad (a)$$

the rate should be proportional to C_A, for

$$2A \longrightarrow \text{Products} \qquad (b)$$
and
$$A + B \longrightarrow \text{Products} \qquad (c)$$
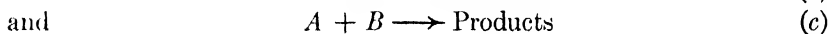
to C_A^2 and $C_A C_B$, respectively, while for

$$A + 2B \longrightarrow \text{Products} \qquad (d)$$
or
$$2A + B \longrightarrow \text{Products} \qquad (e)$$
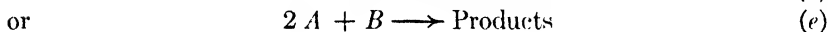
to $C_A C_B^2$ and $C_A^2 C_B$. These proportionalities indicate that different rate dependencies on concentration are obtained at constant temperature for various reactions, and to distinguish these the term *order of reaction* is employed. By the order of a chemical reaction is meant *the sum of all the exponents to which the concentrations in the rate equation are raised.* Thus when the rate of a reaction is given by

$$\frac{- dC}{dt} = k C_1^{n_1} C_2^{n_2} C_3^{n_3} \cdots \qquad (1)$$

where k is a constant, the reaction orders of the individual constituents are n_1, n_2, n_3, etc., and the order of the reaction as a whole, n, is

$$n = n_1 + n_2 + n_3 + \cdots \qquad (2)$$

In examples (a) to (e) the order of reaction is identical with the number of molecules of reactants participating in the reaction. Thus, the first order reaction (a) is also *unimolecular*, the second order reactions (b) and (c) are *bimolecular*, while the third order reactions (d) and (e) are *termolecular*. Although this identity of reaction order and the number of molecules reacting as given by the stoichiometric equation of the

process is observed in many reactions, it does not hold in all cases. For instance, for the reactions

$$CH_3CHO \longrightarrow CH_4 + CO$$

and
$$3\ KClO \longrightarrow KClO_3 + 2\ KCl$$
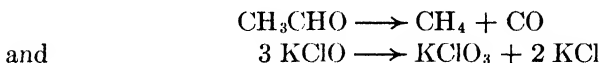
we would anticipate *a priori* that the first of these would be first order, the second third order. Actually both reactions are found to be second order. Since the reaction order *may* differ from the number of molecules participating in the reaction, a distinction must be made between the *molecularity*, i.e., the number of molecules involved in the step leading to reaction, and the order of a reaction. According to the former, reactions are unimolecular, bimolecular, etc., depending on whether one, two, or more molecules are involved in the rate determining step. On the other hand, the term reaction order is confined to the dependence of the observed rate on the concentrations of the reactants.

It is evident that the order, or for that matter the molecularity, of a reaction cannot always be predicted from the stoichiometric equation of a reaction, but that each and every reaction must be investigated kinetically to ascertain both order and molecularity. Further, care must be exercised in interpreting an observed reaction order and in inferring from it the molecularity and the mechanism of the reaction process. When the order of a reaction agrees with the rate equation obtained by applying the law of mass action to the stoichiometric equation of the reaction, it is fairly safe to assume that the reaction proceeds as given by the chemical equation. When the two are not the same, however, a mechanism has to be devised which is in accord both with the observed reaction order and with the overall equation of the process.

In developing the rate equations for reactions of various orders we shall proceed on the supposition that molecularity and order are identical. Whenever necessary attention will be directed to any discrepancy between the two.

First Order Reactions. According to the law of mass action the rate of any unimolecular reaction

$$A \longrightarrow \text{Products} \tag{3}$$

should at any time t be proportional to the concentration of A, C_A, present in the system at that instance. If this proportionality is converted to an equality by insertion of a proportionality constant k_1, we obtain the differential equation for a first order reaction, namely,

$$\frac{-dC_A}{dt} = k_1 C_A \tag{4}$$

The constant k_1 is called the *specific rate* or the *specific velocity constant* of the first order reaction. On setting $C_A = 1$ in equation (4), the significance of this constant may be seen to be the speed of the reaction when the concentration of A is constant and equal to unity. Again, the units of k_1 follow from $k_1 = (-1/C_A)(dC_A/dt)$ as a reciprocal time $1/t$, i.e., a frequency. For any first order reaction k_1 should be a constant characteristic of the reaction, independent of the concentration, and a function of the temperature only.

Before integrating this differential equation, it is desirable to transform it into a more familiar form. To do this let a be the initial concentration of A, and x the decrease in concentration of A due to reaction up to time t. Then $C_A = a - x$ at time t,

$$\frac{-dC_A}{dt} = \frac{-d(a-x)}{dt} = \frac{dx}{dt}$$

and equation (4) becomes

$$\frac{dx}{dt} = k_1(a - x) \tag{5}$$

This equation gives the rate for a first order reaction in terms of the initial concentration and the amount of substance reacted. Integrating equation (5) and remembering that when $t = 0$, i.e., at the start of the reaction, $x = 0$, and that when $t = t \quad x = x$, we obtain:

$$\int_{x=0}^{x=x} \frac{dx}{a-x} = \int_{t=0}^{t=t} k_1 dt$$

$$\left[-\ln(a-x)\right]_{x=0}^{x=x} = \left[k_1 t\right]_{t=0}^{t=t}$$

$$\ln \frac{a}{a-x} = k_1 t \tag{6}$$

Any first order reaction must satisfy equation (6). To ascertain whether a particular reaction obeys this equation, several methods are available. First, knowing the initial concentration and the concentrations of reactant at various elapsed times, a, $(a - x)$, and t corresponding to the latter may be substituted into the equation and k_1 solved for. If the reaction is first order, a series of k_1's are thus obtained for each set of values of $(a - x)$ and t which are constant within the accuracy of the experiment. However, if the k_1 values thus calculated are not constant but exhibit an appreciable drift in magnitude, the reaction is not first order. One of the higher order equations must then be tried to find one that will satisfy the observed data.

An alternate method of testing equation (6) is a graphical one. Equation (6) may be rearranged to the form

$$\ln (a - x) = - k_1 t + \ln a$$

or
$$\log_{10} (a - x) = \left(\frac{- k_1}{2.303}\right) t + \log_{10} a \qquad (7)$$

Since for any one experiment a is constant, equation (7) is the equation of straight line when $\log(a - x)$ is plotted against t. In such a plot the y intercept will be $\log_{10} a$, while the slope will be $(- k_1/2.303)$. Consequently, when a $\log(a - x)$ vs. t plot constructed from the experimental data is found to be linear, the reaction is first order. By taking the slope of the line, k_1 follows as

$$k_1 = - 2.303 \text{ (slope)} \qquad (8)$$

A third means of testing first order reactions is known as the *fractional life method*. In this method the time necessary to decompose a definite fraction of the reactant, usually one-half, is determined for a number of different values of a. When one-half of the reactant has undergone decomposition, $a - x = a/2$, and the time, $t_{1/2}$, necessary for this to occur follows from equation (6) as

$$t_{1/2} = \frac{1}{k_1} \ln (a/a/2)$$
$$= \frac{\ln 2}{k_1} \qquad (9)$$

$t_{1/2}$ is known as the *half-life period* of the reaction. According to equation (9) *the half-life period of any first order reaction is a constant independent of the initial concentration*. In other words, it takes a first order reaction just as much time to go halfway to completion when the initial concentration is high as when it is low. That this requirement is met by first order reactions is shown by the results obtained by Hinshelwood and Hutchison[1] on the thermal decomposition of gaseous acetone at 601° C. These results are reproduced in Table 1. Although the initial pressure of the acetone is seen to vary about fourfold, the half-life period nevertheless remains constant within the experimental accuracy. Taking as an average $t_{1/2} = 81$ sec, we find through equation (9) for the rate constant of this reaction

$$k_1 = \frac{\ln 2}{t_{1/2}} = \frac{0.693}{81}$$
$$= 0.0086 \text{ sec}^{-1}$$

[1] Hinshelwood and Hutchison, Proc. Roy. Soc., **111A**, 245 (1926).

TABLE 1

HALF-LIFE PERIODS FOR THERMAL DECOMPOSITION OF ACETONE AT 601° C

a (mm Hg)	$t_{1/2}$ (sec)
98	86
192	78
230	85
296	80
362	77

In calculating rate constants the common practice is to express the time in minutes or seconds. Occasionally, with very slow reactions the time may be expressed in hours or even days. Concentrations, on the other hand, should invariably be expressed in moles per liter for reactions in solution, and in these units or some suitable pressure units, such as millimeters of mercury or atmospheres, for gaseous reactions. In any case, for complete clarity the units of time and concentration for which k is given should be stated, as well as the temperature. However, for first order reactions, *and only for these*, the units in which the concentrations are expressed do not matter. An inspection of equation (6) will reveal that k_1 depends on the ratio of two concentrations, $a/(a - x)$. Consequently, as long as a and $a - x$ are expressed in the same units the value of the ratio is unaffected, and any convenient units may be employed.

First Order Reactions in Gases. With the exception of the decomposition of nitrogen pentoxide, homogeneous gaseous first order reactions were practically unknown until 1925. Since then quite a few have been discovered and studied, so that at present about 15 or 20 reactions are known which, at least within certain pressure limits, are first order and homogeneous. Among these are included the thermal dissociations of nitrous oxide, nitrogen pentoxide, acetone, propionic aldehyde, various aliphatic ethers, azo compounds, and amines, ethyl bromide, and the isomerization of d-pinene ($C_{10}H_{16}$) to dipentene.

As typical of a homogeneous first order gas reaction may be taken the thermal decomposition of azoisopropane to hexane and nitrogen,

$$\begin{array}{c} CH_3 \\ \diagdown \\ \diagup \\ CH_3 \end{array} CH-N{=}N-CH \begin{array}{c} CH_3 \\ \diagdown \\ \diagup \\ CH_3 \end{array} \longrightarrow N_2 + C_6H_{14}$$
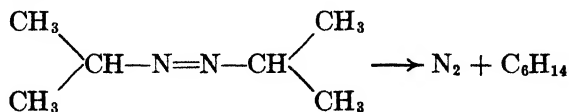
which was investigated by Ramsperger[1] over the pressure range 0.0025 to 46 mm Hg and between 250 and 290° C. The rate of reaction was followed by pressure measurements with a McLeod gage. The only data necessary were the initial pressure of the reactant and the total pressures

[1] Ramsperger, J. Am. Chem. Soc., **50**, 714 (1928).

of the system corresponding to various stages of decomposition. From these pressures the rate constant of the process may be calculated as follows. If we let P_i be the initial pressure of azoisopropane, P the total pressure at any time t, P_A the pressure of azoisopropane at any time t, and x the decrease in pressure of azoisopropane at any time t, then at each stage of the reaction $P_A = P_i - x$, $P_{N_2} = P_{C_6H_{14}} = x$, and the total pressure of the system is given by

$$P = P_A + P_{N_2} + P_{C_6H_{14}}$$
$$= (P_i - x) + x + x$$
$$= P_i + x$$

From this equation we find $x = P - P_i$, and on substitution of this value for x into $P_A = P_i - x$, P_A becomes

$$P_A = P_i - x$$
$$= P_i - (P - P_i)$$
$$= 2 P_i - P$$

Since a in equation (6) is proportional to P_i, and $(a - x)$ to P_A, the expression for k_1 is

$$k_1 = \frac{2.303}{t} \log_{10} \frac{P_i}{P_A}$$
$$= \frac{2.303}{t} \log_{10} \frac{P_i}{2 P_i - P}$$

The constants thus calculated from experimental data for a typical run at 270° C are shown in Table 2. Column 1 gives the time in minutes, column 2 the total pressure P in millimeters of mercury, and column 3 the values of k_1. P_i is the pressure of the system at $t = 0$, which for the experiment in question was 35.15 mm Hg. It is seen that k_1 exhibits satisfactory constancy, and therefore the decomposition of azoisopropane is a first order reaction.

TABLE 2

DECOMPOSITION OF AZOISOPROPANE AT 270° C

t (min)	P (mm Hg)	k_1 (min⁻¹)
0	35.15	—
3	46.30	2.12×10^{-3}
6	53.90	2.11
9	58.85	2.07
12	62.20	2.03
17	65.55	1.96

Mean $= 2.06 \times 10^{-3}$

First Order Reactions in Solution. As an example of a first order reaction in solution may be taken the decomposition of benzene diazo-

nium chloride, C_6H_5N=NCl, in water. This compound, known only in solution, is very unstable, and dissociates readily on heating to liberate nitrogen. As the nitrogen all comes from the compound, the volume of it liberated may be used as a measure of the decrease in concentration of the diazo salt. Cain and Nicoll[1] utilized this fact in their kinetic study of this reaction. A freshly prepared solution of the salt was quickly heated to the desired temperature, and the reaction was then permitted to proceed in a thermostat. The course of the reaction was followed by measuring in a gas burette the volumes V of nitrogen evolved at various time intervals. To complete the required data, the initial concentration expressed as cubic centimeters of nitrogen V_0 was calculated from the amount of diazo compound originally present. In terms of V_0 and V, the velocity constant of the reaction follows as

$$k_1 = \frac{2.303}{t} \log_{10}\left(\frac{a}{a-x}\right) = \frac{2.303}{t} \log_{10}\left(\frac{V_0}{V_0 - V}\right)$$

The results obtained in an experiment at 50° C with an amount of benzene diazonium chloride equivalent to 58.3 cc nitrogen are shown in Fig. 2, where a plot of $\log_{10}(a - x) = \log_{10}(58.3 - V)$ vs. t is given.
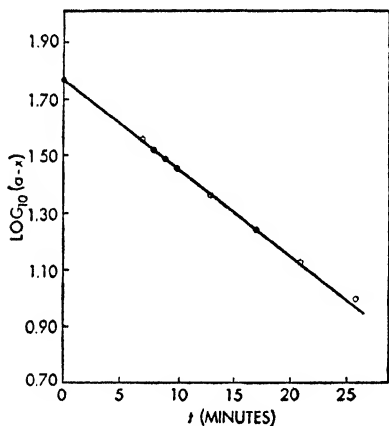
Fig. 2. Plot of $\text{Log}_{10}(a - x)$ vs. t for First Order Reactions

As is to be anticipated for a first order reaction, the plot is linear. Taking the slope of this line we find slope $= -0.0308$, and consequently k_1, according to equation (8), is

$$k_1 = -2.303 \text{ (slope)}$$
$$= -2.303 \ (-0.0308)$$
$$= 0.071$$

As examples of other first order reactions in solution may be mentioned the decomposition of nitrogen pentoxide and the isomerization of d-pinene in various organic solvents, the decomposition of malonic, trichloracetic, and acetonedicarboxylic acids in water, and the dissociation of various diazonium salts in the same solvent.

Second Order Reactions. Any bimolecular reaction can in general be represented by the equation

$$A + B \longrightarrow \text{Products} \tag{10}$$

[1] Cain and Nicoll, J. Chem. Soc., **81,** 1412 (1902).

If we let a and b be the initial concentrations of A and B respectively, and let x be the decrease in concentration of each at time t, then at t the concentration of A will be $(a - x)$, that of B $(b - x)$, and the rate equation for the reaction becomes

$$\frac{dx}{dt} = k_2(a - x)(b - x) \tag{11}$$

Equation (11) is the rate equation for a *second order* reaction. k_2 is the specific rate constant for such a reaction, and is equal to the rate when both A and B are unity. Its units are $1/tC$, or $t^{-1}C^{-1}$, and its magnitude depends on the nature of the reaction, the temperature, and the concentration and time units in which it is expressed. Integration of this equation, again remembering that $x = 0$ when $t = 0$, yields, for k_2,

$$k_2 = \frac{1}{t(a - b)} \ln \frac{b(a - x)}{a(b - x)}$$

or
$$k_2 = \frac{2.303}{t(a - b)} \log_{10} \frac{b(a - x)}{a(b - x)} \tag{12}$$

This is the equation which any second order reaction must obey.

Equation (12) is considerably simplified when A and B are present in equal initial concentrations, or when A and B are the same, as in the reaction

$$2A \longrightarrow \text{Products} \tag{13}$$

For these conditions equation (11) reduces to

$$\frac{dx}{dt} = k_2(a - x)^2 \tag{14}$$

and on integration becomes either

$$\left(\frac{1}{a - x}\right) = k_2 t + \frac{1}{a} \tag{15}$$

or
$$k_2 = \frac{1}{at}\left(\frac{x}{a - x}\right) \tag{16}$$

which are identical.

These rate expressions for second order reactions may be tested by the methods described for first order reactions. Values of a, b, and x, or a and x may be substituted along with t and k_2 solved for, or k_2 may be evaluated graphically. For the latter procedure equation (12) is best written in the form

$$t = \frac{2.303}{k_2(a - b)} \log_{10} \frac{(a - x)}{(b - x)} + \frac{2.303}{k_2(a - b)} \log_{10}\left(\frac{b}{a}\right) \quad \cdot \tag{17}$$

from which we see that, since k_2, a, and b are constant for any experiment, a plot of t against $\log_{10}(a - x)/(b - x)$ should yield a straight line with

$$\text{Slope} = \frac{2.303}{k_2(a - b)} \tag{18}$$

and y intercept equal to $\dfrac{2.303}{k_2(a - b)} \log_{10}\left(\dfrac{b}{a}\right)$. By taking the slope of such a plot and substituting into equation (18), k_2 can readily be obtained. On the other hand, to test the rate law when the concentrations are the same, equation (15) is used. According to this equation a plot of $1/(a - x)$ vs. t should be linear, with slope equal to k_2 and the y intercept equal to $1/a$. By taking the slope k_2 is obtained directly.

The half-life method cannot be used with reactions where the concentrations of A and B are different, since A and B will have different times for half reaction. However, it may be used where the concentrations are the same, or where the two reacting molecules are identical. In these instances $x = a - x = a/2$ at the half-life point, and the period of half-life, $t_{1/2}$, follows from equation (16) as

$$t_{1/2} = \frac{1}{k_2} \cdot \frac{\left(\dfrac{a}{2}\right)}{a\left(\dfrac{a}{2}\right)}$$

$$= \frac{1}{k_2 a} \tag{19}$$

i.e., *for a second order reaction the period of half-life is inversely proportional to the first power of the initial concentration.* Knowing $t_{1/2}$ and a, k_2 is readily evaluated from equation (19).

It may be stated as a general rule that the period of half-life for any n order reaction is inversely proportional, but not equal, to the $(n - 1)$ power of the initial concentration, or

$$t_{1/2} \propto \frac{1}{a^{n-1}} \tag{20}$$

Thus, for $n = 1$ $t_{1/2}$ is independent of a, for $n = 2$ inversely proportional to it, for $n = 3$ inversely proportional to a^2, etc. The proportionality constant for equation (20) depends on the reaction order, and is $\ln 2/k_1$ for first order, $1/k_2$ for second order, and $3/2\, k_3$ for third order reactions.

Second Order Gas Reactions. Of all homogeneous gas reactions those of second order are encountered most frequently. In this category are included various thermal dissociations, such as those of hydrogen iodide, nitrogen dioxide, ozone, chlorine monoxide, nitrosyl chloride,

formaldehyde, and acetaldehyde; the combination of hydrogen and iodine to form hydrogen iodide; the polymerization of ethylene; the hydrogenation of ethylene; and others.

To illustrate the behavior of all such reactions, the thermal decomposition of acetaldehyde may be taken. This reaction, when investigated by Hinshelwood and Hutchison,[1] was found to be almost totally homogeneous and to proceed smoothly according to the equation

$$2 \, CH_3CHO \longrightarrow 2 \, CH_4 + 2 \, CO$$
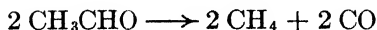
Since in this reaction there is an increase in pressure at constant volume on dissociation, the change in pressure observed on a manometer attached to the system may be employed to follow the reaction course. From these pressure measurements k_2 is calculated as follows.

If we let P_i be the initial pressure of acetaldehyde, and x the decrease in pressure after time t, then the pressure of the reactant at time t is $(P_i - x)$. The stoichiometry of the reaction indicates that, when the pressure of acetaldehyde drops by x, the pressures of methane and carbon monoxide must each increase by x. We obtain thus for the total pressure of the system P, at any time t,

$$\begin{aligned} P &= P_{CH_3CHO} + P_{CH_4} + P_{CO} \\ &= (P_i - x) + x + x \\ &= P_i + x \end{aligned}$$

from which x follows as

$$x = P - P_i$$

Therefore, by measuring the initial pressure and the total pressure of the system at various time intervals x may be evaluated. Further, since $P_i - x$ is proportional to $(a - x)$ and P_i to a, substitution of these into equation (16) yields for k_2 the expression

$$\begin{aligned} k_2 &= \frac{1}{at}\left(\frac{x}{a - x}\right) \\ &= \frac{1}{P_i t}\left(\frac{x}{P_i - x}\right) \end{aligned}$$

In Table 3 are given data obtained by Hinshelwood and Hutchison for this reaction during an experiment at 518° C with an initial pressure of acetaldehyde equal to 363 mm Hg. Column 1 gives the elapsed time in seconds, column 2 $x = P - P_i$ in mm Hg, and column 3 the values of k_2 calculated from these in the manner described. The good constancy of the latter confirms both the reaction order and the correctness of the formulation.

[1] Hinshelwood and Hutchison, Proc. Roy. Soc., 111A, 380 (1926).

TABLE 3

THERMAL DECOMPOSITION OF ACETALDEHYDE AT 518° C

$(P_i = 363 \text{ mm Hg})$

t (sec)	$x = P - P_i$ (mm Hg)	$k_2 = \dfrac{1}{P_i t}\left(\dfrac{x}{P_i - x}\right)$
42	34	6.79×10^{-6}
73	54	6.59
105	74	6.71
190	114	6.64
242	134	6.66
310	154	6.55
384	174	6.61
480	194	6.59
665	224	6.68
840	244	6.72
1070	264	6.86
1440	284	6.88

Average $= \overline{6.69 \times 10^{-6}}$

Second Order Reactions in Solution. Among the many reactions in solution which are of second order may be mentioned the saponifications of various esters by bases, the conversion of nitroparaffins to aci-nitroparaffins by bases, the reactions of alkyl halides with amines, the hydrolyses of esters, amides, and acetals, esterifications of organic acids, the combination of NH_4^+ and CNO^- ions to form urea, and the reactions of bromacetates with thiosulfates and thiocyanates.

As a prototype of these reactions may be taken the saponification of ethyl butyrate by hydroxyl ions in water solution, namely,

$$CH_3CH_2CH_2COOC_2H_5 + OH^- \longrightarrow CH_3CH_2CH_2COO^- + C_2H_5OH$$
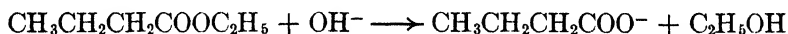
which was studied by Williams and Sudborough[1] at 20° C. The reaction mixture was prepared by bringing together solutions of ethyl butyrate and barium hydroxide so as to yield an initial concentration a of the ester and b of the base. The course of the reaction was followed by removing periodically samples of solution and titrating these with standard acid to ascertain the concentration of unreacted barium hydroxide. The concentrations of base thus determined give directly $(b - x)$, while a minus the amount of reacted base yields $(a - x)$. Table 4 shows

TABLE 4

SAPONIFICATION OF ETHYL BUTYRATE BY BARIUM HYDROXIDE AT 20° C

t (min)	$a - x$ (arbitrary units)	$b - x$	k_2 [equation (12)]
0	19.75	20.85	—
5	14.75	15.85	0.0032
17	9.40	10.50	0.0030
36	5.93	7.03	0.0029
65	3.57	4.67	0.0030

[1] Williams and Sudborough, J. Chem. Soc., **101**, 415 (1912).

some of the data obtained in this manner, $(a - x)$ and $(b - x)$ being given in arbitrary but convenient units, and the value of k_2 calculated from these by equation (12). As an alternate test, t may be plotted against $\log_{10}(a - x)/(b - x)$ to see whether a straight line results in accordance with equation (17). Such a plot, given in Fig. 3, is linear.

Third Order Reactions. The most general type of termolecular reaction is one where three different molecules react, namely,

$$A + B + C \longrightarrow \text{Products} \quad (21)$$
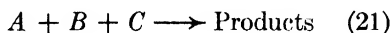
If the reactants are all present at different initial concentrations a, b, c, and x is the decrease in the concentration of each at time t, then the *third order* rate equation is

$$\frac{dx}{dt} = k_3(a - x)(b - x)(c - x) \quad (22)$$
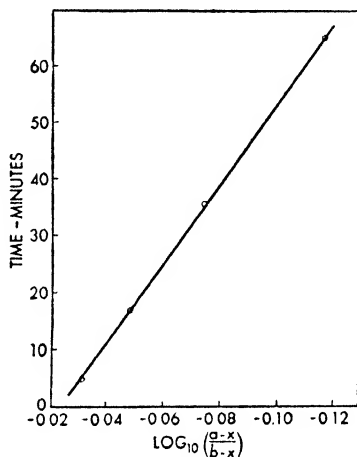
Fig. 3. Graphical Test of Second Order Reactions

The integrated form of equation (22) is rather complex. For a simpler case where two of the initial concentrations are equal, say $a = b$, equation (22) reduces to

$$\frac{dx}{dt} = k_3(a - x)^2(c - x) \quad (23)$$

which on integration yields, for k_3,

$$k_3 = \frac{1}{t(c - a)^2}\left[\frac{x(c - a)}{a(a - x)} + \ln\frac{c(a - x)}{a(c - x)}\right] \quad (24)$$

On the other hand, when two of the molecules are identical, as in

$$2\,A + B \longrightarrow \text{Products} \quad (25)$$

the concentrations at any time are $(a - 2\,x)$, $(b - x)$, the rate equation becomes

$$\frac{dx}{dt} = k_3(a - 2\,x)^2(b - x) \quad (26)$$

and k_3 is given by

$$k_3 = \frac{1}{t(2\,b - a)^2}\left[\frac{2\,x(2\,b - a)}{a(a - 2\,x)} + \ln\frac{b(a - 2\,x)}{a(b - x)}\right] \quad (27)$$

The simplest case, of course, is a reaction of the type

$$3\,A \longrightarrow \text{Products} \tag{28}$$

or reaction (21) with all concentrations equal. Then the third order rate equation is

$$\frac{dx}{dt} = k_3(a - x)^3 \tag{29}$$

and, on integration,

$$k_3 = \frac{1}{2\,t}\left[\frac{1}{(a-x)^2} - \frac{1}{a^2}\right] \tag{30}$$

$$= \frac{1}{2\,ta^2}\left[\frac{x(2\,a - x)}{(a-x)^2}\right] \tag{31}$$

From equation (31) the period of half-life of a third order reaction follows readily, on setting $x = a/2$, as

$$t_{1/2} = \frac{3}{2\,k_3 a^2} \tag{32}$$

i.e., for such a reaction $t_{1/2}$ is inversely proportional to a^2.

Only five homogeneous gas reactions are definitely known to be third order, and everyone of these involves the interaction of nitric oxide (NO) with either chlorine, bromine, oxygen, hydrogen, or heavy hydrogen. As an example of one of these may be taken the work of Hinshelwood and Green [1] on the reaction

$$2\,NO + H_2 \longrightarrow N_2O + H_2O$$
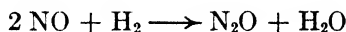
For hydrogen present in excess, k_3 for this reaction should be given by equation (27). Now, at a time t' when the nitric oxide is half consumed $2\,x = a/2$, and this equation becomes

$$k_3 = \frac{1}{t'(2\,b - a)^2}\left[\frac{2\,b - a}{a} + \ln\frac{2\,b}{4\,b - a}\right] \tag{33}$$

This is the relation Hinshelwood and Green used to evaluate k_3 from various initial pressures of nitric oxide ($a = P_{NO}$) and hydrogen ($b = P_{H_2}$), and the observed times, t', of half decomposition of nitric oxide. Their results, given in Table 5, yield fairly satisfactory constants, and indicate that the reduction of nitric oxide by hydrogen is a third order reaction.

[1] Hinshelwood and Green, J. Chem. Soc., **128**, 730 (1926).

<div align="center">

TABLE 5

KINETICS OF THE REACTION

$2 NO + H_2 \longrightarrow N_2O + H_2O$ at 826° C

</div>

P_{NO} (mm Hg)	P_{H_2} (mm Hg)	t' (sec)	k_3
110	316	270	1.19×10^{-7}
152	404	204	0.91
359	400	89	1.12
144	323.5	227	1.10
181	209.5	264	1.39
370	376	92	1.17
300	404	100	1.09
232	313	152	1.19

Reactions in solution which appear to be third order are the oxidation of ferrous sulfate in water, the reaction between iodide and ferric ions in aqueous solution, and the action between benzoyl chloride and alcohols in ether solution.[1] The decomposition of hypobromous acid at constant pH in the pH range of 6.4 to 7.8 has also been found to be third order with respect to the acid.[2]

Pseudo-Molecular Reactions. There are many reactions which obey a first order rate equation although in reality they are bi- or termolecular. As an example of these may be taken the decomposition of carbonyl sulfide in water, namely,

$$COS + H_2O \longrightarrow CO_2 + H_2S$$

According to the law of mass action this reaction should be second order, with the rate dependent on the concentration of both the carbonyl sulfide and the water. Actually, however, the rate is found to be first order with respect to the carbonyl sulfide and independent of the water. Reactions exhibiting such behavior are said to be *pseudo-unimolecular*.

The pseudo-unimolecular nature of all reactions of this type is explainable by the fact that the concentration of one of the reactants, in this case water, is present in such excess that its concentration remains practically constant during the course of the reaction. Under these conditions $b - x = b$, and the rate equation becomes

$$\frac{dx}{dt} = k_2(a - x)b \tag{34}$$

On integration this leads to

$$k = bk_2 = \frac{2.303}{t} \log_{10} \frac{a}{a - x} \tag{35}$$

which is the equation for a first order reaction.

[1] Moelwyn-Hughes, "Kinetics of Reactions in Solution," Oxford University Press, Oxford, 1933, p. 258.

[2] Prutton and Maron, J. Am. Chem. Soc., **57**, 1652 (1935).

It is evident, however, that the new constant k is not independent of the concentration, as is the case with true first order constants, but may vary with the concentration of b as the latter is changed appreciably. When such is the case the true constant k_2 can be obtained from k by dividing the latter by b, i.e., $k_2 = k/b$.

Pseudo-molecular reactions are encountered whenever one or more of the reactants remain constant during the course of an experiment. This, for one, is the case with reactions conducted in solvents which are themselves one of the reactants, as in the decomposition of carbonyl sulfide in water, or in the esterification of acetic anhydride in alcohol:

$$(CH_3CO)_2O + 2\ C_2H_5OH \longrightarrow 2\ CH_3COOC_2H_5 + H_2O$$
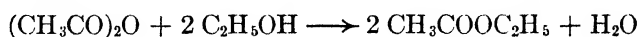
Again, this is true of reactions subject to catalysis in which the concentration of the catalyst does not change. The decomposition of diacetone alcohol to acetone in aqueous solution is catalyzed by hydroxyl ions, with the rate proportional to the concentration of the alcohol and that of the base. Since the concentration of the base does not change within any one experiment, however, the rate equation reduces to one of first order with respect to the alcohol. But, the rate constants obtained for various concentrations of base are not identical; in fact, they increase with the latter, as may be seen from column 2 of Table 6. To obtain from these pseudo-unimolecular k's the true second order velocity constant, the k's must be divided by the hydroxyl ion concentration. When this is done excellent k_2 values result, as column 3 indicates.

<div align="center">

TABLE 6

DECOMPOSITION OF DIACETONE ALCOHOL IN WATER AT 25° C [1]

(Catalyst: NaOH)

</div>

C Conc. of NaOH	k (min^{-1})	$k_2 = k/C$
0.0205	0.0455	2.22
0.0292	0.0651	2.23
0.0518	0.1160	2.24
0.0710	0.1576	2.22
0.1045	0.2309	2.21

Reversible or Opposing Reactions. In formulating the rate equations for the reactions considered so far, the tacit assumption has been made that the processes proceed in the direction indicated without any tendency for the reactions to reverse themselves. Such a treatment is justifiable only with reactions whose point of equilibrium lies way over on the side of the products. In other reactions the products formed set in a reaction counter to the forward process whose rate increases as the products accumulate, and eventually becomes equal to that of the for-

[1] LaMer and Miller, J. Am. Chem. Soc., **57**, 2674 (1935).

ward reaction. At this point the overall rate becomes zero, and the system is in equilibrium.

Reactions exhibiting this tendency to reverse themselves are called *reversible* or *opposing reactions*. As an illustration of one of these and of the method of dealing with them kinetically may be taken the oxidation of nitric oxide to nitrogen dioxide,

$$2\,NO + O_2 \underset{k_2}{\overset{k_3}{\rightleftharpoons}} 2\,NO_2$$
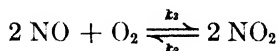
investigated by Bodenstein and Lindner.[1] These authors found that below 290° C the rate of the forward reaction is third order and proceeds as written without complications. However, above this temperature the rate of dissociation becomes noticeable and leads to a decrease in the rate of disappearance of nitric oxide and oxygen. To correct for this back reaction we proceed as follows. If we let a and b be the initial concentrations of nitric oxide and oxygen respectively, and x the amount of oxygen reacted in time t, then at t the concentration of oxygen is $(b - x)$, of nitric oxide $(a - 2\,x)$, while that of nitrogen dioxide formed $2\,x$. The rate of the forward reaction is hence given by

$$\left(\frac{dx}{dt}\right)_f = k_3(a - 2\,x)^2(b - x) \tag{36}$$

that of the reverse reaction by

$$\left(\frac{dx}{dt}\right)_r = k_2(2\,x)^2 \tag{37}$$

Since these two rates oppose each other, the net rate of the forward reaction, dx/dt, must be the difference between them, or

$$\begin{aligned}
\frac{dx}{dt} &= \left(\frac{dx}{dt}\right)_f - \left(\frac{dx}{dt}\right)_r \\
&= k_3(a - 2\,x)^2(b - x) - k_2(2\,x)^2
\end{aligned} \tag{38}$$

This equation alone is insufficient to evaluate both k_3 and k_2 from experimental data. To do this another relation between k_3 and k_2 is required. For obtaining this relation we utilize the fact that when the reaction reaches equilibrium $dx/dt = 0$, and therefore

$$k_2 = \frac{k_3(a - 2\,x_e)^2(b - x_e)}{(2\,x_e)^2} \tag{39}$$

where x_e is the value of x at equilibrium. On substitution of this value of k_2 into the integrated form of equation (38), an expression is obtained involving only k_3 which may be used to test the kinetics of the reaction.

[1] Bodenstein and Lindner, Z. physik. Chem., **100**, 87 (1922).

It will be observed that here not only the initial concentrations and x at various times are required, but also the value of x at equilibrium, x_e. To find the latter, the reaction is permitted to proceed until equilibrium is attained, and x_e is measured.

In Table 7 are shown some results obtained by Bodenstein and Lindner during an experiment at 339° C. Column 1 gives the time in minutes, column 2 the values of k_3 calculated by using only equation (36), and column 3 the values of k_3 obtained after correction for the back reaction by means of equation (38). Whereas the constants in column 2 diminish regularly with time, those in column 3 are constant, and confirm the applicability of equation (38) to this reaction.

Other reactions which are reversible, and which can be treated in a manner analogous to the one described are the mutarotation of glucose, the combination of hydrogen and iodine to form hydrogen iodide, the hydrogenation of ethylene, and the hydrolyses of various esters.

TABLE 7

KINETICS OF THE REACTION

$2 NO + O_2 \longrightarrow 2 NO_2$ at 339° C

t (min)	k_3 [equation (36)]	k_3 [equation (38)]
4	0.58×10^{-5}	0.58×10^{-5}
5	0.58	0.59
6	0.57	0.59
8	0.54	0.58
10	0.53	0.59
12	0.50	0.59
17	0.46	0.61
22	0.37	0.63

Consecutive Reactions. Chemical reactions, such as

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \qquad (40)$$

which proceed from reactants to products through one or more intermediate stages are called *consecutive reactions*. In these reactions each stage has its own rate and its own velocity constant. Further, whether one or another of these particular rates, or some combination of all, is measured experimentally depends on the relative magnitudes of the velocity constants of the various stages. It is a well-established fact that in any sequence of reactions of varying speed the one that will determine the rate of the overall reaction is the one that proceeds with the *slowest* speed. This follows of necessity because any succeeding stage has to wait for any preceding before it can take place. Consequently, if, in a reaction such as that represented by equation (40), k_1 is very much greater than k_2, the conversion of B to C will determine the rate of formation of the product. On the other hand, if k_2 is very much greater than

k_1, the formation of B from A will control the rate, and C will result from B as soon as the latter appears. However, when k_1 and k_2 are comparable in magnitude, the rate of the overall reaction depends on both constants, and the situation becomes more complex.

As an instance of a consecutive reaction proceeding in two stages with appreciably different velocity constants may be cited the decomposition of sodium hypochlorite in aqueous alkaline solutions. Although the stoichiometric equation of the process,

$$3 \text{ NaClO} \longrightarrow 2 \text{ NaCl} + \text{NaClO}_3 \qquad (41)$$

suggests a third order reaction, the process is actually second order. To explain this result the suggestion has been advanced that the reaction proceeds in two steps, the first of which involves two molecules of sodium hypochlorite to form sodium chlorite,

$$2 \text{ NaClO} \longrightarrow \text{NaCl} + \text{NaClO}_2 \qquad (41a)$$

and the second a molecule of sodium chlorite and sodium hypochlorite to yield sodium chlorate,

$$\text{NaClO}_2 + \text{NaClO} \longrightarrow \text{NaCl} + \text{NaClO}_3 \qquad (41b)$$

The sum of equations (41a) and (41b) is, of course, equation (41). To decide which of these two reactions is rate determining, Förster and Dolch[1] investigated not only the decomposition of sodium hypochlorite but also the reaction between sodium chlorite and hypochlorite to form the chlorate. They found that the rate of the latter reaction is about 25 times that of the former, and consequently the rate of equation (41a) must control the rate of the overall reaction given by equation (41).

The exact mathematical analysis of consecutive reactions with comparable velocity constants is as a rule very difficult unless the reactions are of the simplest kind, such as the one given in equation (40). To show the method of attack, let a be the initial concentration of A, x the amount of it decomposed in time t, and y the concentration of C formed at the same instant. Then at time t we have, $C_A = (a - x)$, $C_C = y$, $C_B = x - y$, since the total amount of B formed is x, and y if it has decomposed to form C. With these concentrations the rate of disappearance of A follows as

$$\frac{dx}{dt} = k_1(a - x) \qquad (42)$$

while the rate of decomposition of B, which is the same as the rate of formation of C, namely, dy/dt, is

$$\frac{dy}{dt} = k_2(x - y) \qquad (43)$$

[1] Förster and Dolch, Z. Elektrochem., **23**, 137 (1917).

To find the dependence of the concentrations of A, B, and C on time these rate equations must first be solved. This is accomplished by integrating equation (42) first, finding x, substituting it into equation (43), and then integrating the latter. As a result we obtain the following expressions for the concentrations:

$$C_A = (a - x) = ae^{-k_1 t} \tag{44a}$$

$$C_B = (x - y) = \frac{k_1 a}{k_2 - k_1} \left(e^{-k_1 t} - e^{-k_2 t}\right) \tag{44b}$$

$$C_C = \frac{a}{k_2 - k_1} \left[(k_2 - k_2 e^{-k_1 t}) - (k_1 - k_1 e^{-k_2 t})\right] \tag{44c}$$

Figure 4 shows C_A, C_B, and C_C as a function of time for $a = 1$, $k_1 = 0.01$, and $k_2 = 0.02$. It will be observed that, whereas the concentration of A falls continuously and that of C increases continuously, the concentration of B rises to a maximum and then falls with time. This behavior of the concentration of the intermediate product is characteristic of consecutive reactions with comparable values of velocity constants, and may be used to identify such reactions.
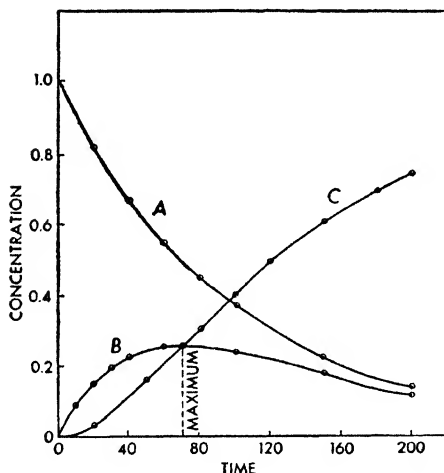
Fig. 4. Dependence of Concentration on Time in Consecutive Reactions

In practice the picture is not quite so simple because as a rule the overall reaction is known and the intermediate steps have to be deduced from it. Under these conditions it may be necessary to follow not only the rate of disappearance of A, but also the rate of formation of B and C before a complete analysis of the reaction may be possible. Further, with reactions of order higher than the first the mathematics become much more formidable, and various artifices have to be relied upon to arrive at a solution of the kinetic processes. Nevertheless, a number of consecutive reactions have been unraveled, and these substantiate the general validity of the attack outlined above.

Side Reactions. *Side reactions* constitute still a third type of complication in kinetic studies. In these reactions the reacting substances, instead of proceeding along one path to yield a given set of products,

follow also one or more other paths to give different products. Thus, in the nitration of phenol, p- and o-nitrophenols are formed simultaneously from phenol and nitric acid by the parallel reactions:

$$\text{OH} \quad + HNO_3 \xrightarrow{k_1} \text{OH}NO_2 + H_2O \tag{45a}$$

o-nitrophenol

$$\text{OH} \quad + HNO_3 \xrightarrow{k_2} \text{OH} + H_2O \tag{45b}$$
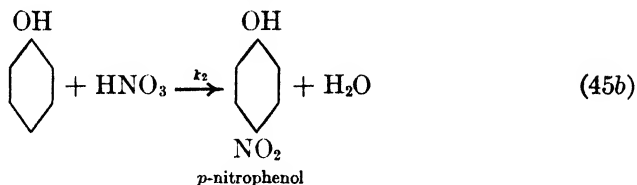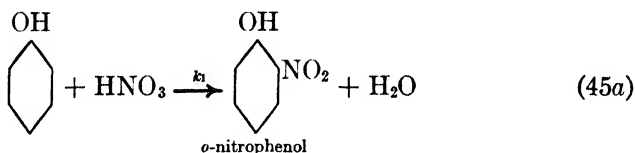
p-nitrophenol

If we let a and b be the initial concentration of phenol and nitric acid, and x the amount of these which react, then at any instant the rate of formation of o-nitrophenol would be given by

$$\frac{d(o\text{-nitrophenol})}{dt} = k_1(a - x)(b - x) \tag{46}$$

and that of p-nitrophenol by

$$\frac{d(p\text{-nitrophenol})}{dt} = k_2(a - x)(b - x) \tag{47}$$

From these the rate of disappearance of the reactants, dx/dt, follows as the *sum* of equations (46) and (47), or

$$\frac{dx}{dt} = (k_1 + k_2)(a - x)(b - x) \tag{48}$$

Again, the ratio of equation (46) to equation (47) yields

$$\frac{d(o\text{-nitrophenol})}{d(p\text{-nitrophenol})} = \frac{k_1(a - x)(b - x)}{k_2(a - x)(b - x)}$$

$$= \frac{k_1}{k_2} \tag{49}$$

Consequently, by following the rate of disappearance of the reactants the sum of the rate constants can be found, while from the rates of formation of the individual products the ratio k_1/k_2 can be obtained. From this sum and ratio the values of the individual constants can be evaluated.

Side reactions are very common, particularly in organic chemistry. The reaction yielding the largest amount of a product is generally re-

ferred to as the *main reaction,* while the others, yielding smaller amounts of products, are called the *side reactions.*

Effect of Temperature on Reaction Velocity. Attention has already been directed to the fact that increase in temperature leads to a tremendous increase in reaction velocity and hence in rate constants. The only known exception to the generality of this statement is the formation of nitrogen dioxide, namely, $2 NO + O_2 \longrightarrow 2 NO_2$, which exhibits a small negative temperature coefficient. This rapid acceleration of reaction rate is observed both in gaseous and liquid phase reactions, as may be seen from the typical data cited in Table 8. In the second order decomposition of gaseous acetaldehyde an increase of $162°$ C in the temperature results in a 450-fold increase in the velocity constant. On the other hand, in the first order decomposition of acetonedicarboxylic acid,

$$CO(CH_2COOH)_2 \longrightarrow CO(CH_3)_2 + 2 CO_2$$

in aqueous solution the constants increase 2200-fold for only a $60°$ C change in temperature. For the latter reaction are tabulated also the half-life periods at various temperatures. These exhibit the tremendous range of half decomposition times of 28,200 to 12 sec on passing from $0°$ to $60°$ C!

Arrhenius first pointed out that the variation of rate constants with temperature can be represented by an equation similar to that used for equilibrium constants, namely,

$$\frac{d \ln k}{dT} = \frac{\Delta E^*}{RT^2} \tag{50}$$

TABLE 8

VARIATION OF REACTION RATE CONSTANTS WITH TEMPERATURE

Decomposition of CH_3CHO in Gas Phase *		Decomposition of $CO(CH_2COOH)_2$ in Aqueous Solution †		
$T°$ K	k_2 (sec^{-1}, moles/liter^{-1})	$T°$ K	k_1 (sec^{-1})	$t_{1/2}$ (sec)
703	0.011	273	2.46×10^{-5}	28,200.
733	0.035	283	10.8	6,410.
759	0.105	293	47.5	1,460.
791	0.343	303	163.	425.
811	0.79	313	576.	120.
836	2.14	323	1850.	37.4
865	4.95	333	5480.	12.6

* Hinshelwood and Hutchison, Proc. Roy. Soc., **111A**, 380 (1926).
† Wiig, J. Phys. Chem., **34**, 596 (1930).

In this *Arrhenius equation* k is the reaction rate constant, T the *absolute* temperature, R the gas constant in calories, and ΔE^* a quantity characteristic of the reaction with the dimensions of an energy. ΔE^* is known as the *energy of activation*, and plays a very important role in chemical kinetics. Its significance will be discussed presently.

If equation (50) is integrated on the supposition that ΔE^* is a constant, we obtain

$$\ln k = -\frac{\Delta E^*}{RT} + C'$$

or

$$\log_{10} k = \left(\frac{-\Delta E^*}{2.303\,R}\right)\frac{1}{T} + C \tag{51}$$

where C' and C are constants of integration. However, if the integration is carried out between the limits $k = k_1$ at $T = T_1$ and $k = k_2$ at $T = T_2$, then,

$$\log_{10}\frac{k_2}{k_1} = \frac{\Delta E^*}{2.303\,R}\left(\frac{T_2 - T_1}{T_1 T_2}\right) \tag{52}$$

From equation (52) it is evident that as soon as two values of k are available at two different temperatures ΔE^* may be evaluated; or, when ΔE^* and a value of k at some one temperature are known, k at another temperature may be calculated. Thus, for the decomposition of acetonedicarboxylic acid, $k = 2.46 \times 10^{-5}$ at $273°$ K and 163×10^{-5} at $303°$ K. Substituting these values into equation (52) we find for ΔE^*,

$$\begin{aligned}
\Delta E^* &= \frac{2.303\,RT_1 T_2}{T_2 - T_1}\log_{10}\frac{k_2}{k_1} \\
&= \frac{2.303 \times 1.987 \times 273 \times 303}{303 - 273}\log_{10}\frac{163 \times 10^{-5}}{2.46 \times 10^{-5}} \\
&= 23{,}000 \text{ cal}
\end{aligned}$$

This value of ΔE^* may now be employed to estimate the values of k's at other temperatures.

To test the validity of the Arrhenius equation we resort to equation (51). According to this equation a plot of $\log_{10} k$ against $1/T$ should be a straight line with

$$\text{slope} = \frac{-\Delta E^*}{2.303\,R} = \frac{-\Delta E^*}{4.58} \tag{53}$$

and y intercept equal to C. Consequently, if such a plot is constructed and is found to be linear, the equation is confirmed. Further, by taking the slope of the line, ΔE^* may be calculated readily through equation (53).

In Fig. 5 are shown plots of $\log_{10} k$ vs. $1/T$ for the data given in Table 8. As predicted by the Arrhenius equation, both reactions yield satisfactory straight lines. From these we find the slope for the acetaldehyde decomposition curve to be -9920, while that for the acetonedicarboxylic acid decomposition -5070. Hence the energy of activation for the first of these reactions is

$$\Delta E^* = -4.58(-9920)$$
$$= 45,500 \text{ cal}$$

and for the second,

$$\Delta E^* = -4.58(-5070)$$
$$= 23,200 \text{ cal}$$
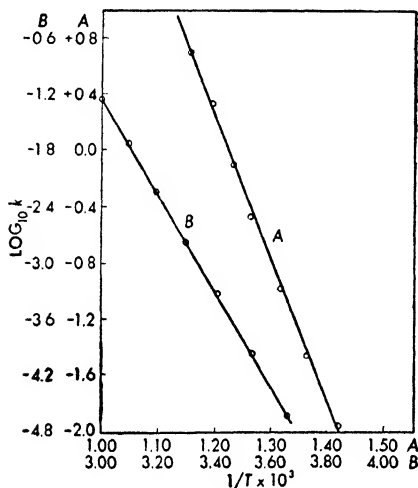
Fig. 5. Variation of Reaction Rate with Temperature. *A*, Decomposition of Acetaldehyde in Gas Phase; *B*, Decomposition of Acetonedicarboxylic Acid in Aqueous Solution

The Energy of Activation. In reactions involving two or more molecules it is logical to presuppose that before reaction can take place the molecules must come into contact with each other; in other words, they must collide. If collision is a sufficient cause for reaction, the rate of reaction should equal the rate of collision. However, when the number of molecules actually reacting in a gaseous reaction, as obtained from the observed velocity constants, is compared with the total number of molecules colliding, calculated from kinetic theory, the latter are found to exceed the number undergoing transformation by many powers of ten. This discrepancy can be explained only by the assumption that molecules in order to react must be in some special configuration at collision, or that they must be in some exceptionally high energy state, or both. Although configuration does play a part in certain reactions, the appearance of the ΔE^* term in the Arrhenius equation and other considerations definitely favor an exceptional energy· state as a prime requisite for reaction; i.e., *molecules must be activated before they can react on collision.*

According to the concept of activation, reactants do not pass directly to products, but must first acquire sufficient energy to pass over an activation "energy barrier." The ideas involved can be made clear with the aid of Fig. 6. In this figure *A* represents the average energy of the reactants, *C* that of the products, and *B* the minimum energy which the reactants must possess in order to react. Molecules existing in state *B*

are said to be *activated* or to be in an *activated state*. Since molecules must be activated before reaction, the reaction must proceed from A to C, not directly, but along the path ABC. In other words, molecules must first climb the energy barrier before they can roll down the hill to form products.

The energy that the reactants at A must absorb in order to become activated and react is the energy of activation, ΔE_1^*, of the process $A \longrightarrow C$. This energy is obviously $\Delta E_1^* = E_B - E_A$, i.e., the difference in energy between the activated state and that corresponding to the average energy of the reactants. By the same argument we see that the energy of activation, ΔE_2^*, of the reverse process $C \longrightarrow A$ must be $\Delta E_2^* = E_B - E_C$. From these two activation energies the difference, ΔE, between ΔE_1^* and ΔE_2^* follows as
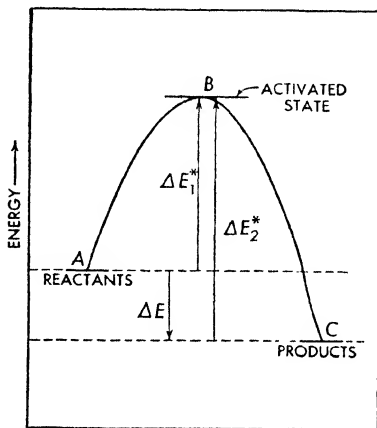
Fig. 6. Energy of Activation

$$\begin{aligned} \Delta E &= \Delta E_1^* - \Delta E_2^* \\ &= (E_B - E_A) - (E_B - E_C) \\ &= E_C - E_A \end{aligned} \tag{54}$$

But $\Delta E = E_C - E_A$ is the difference in energy between reactants and products, or the *heat of reaction at constant volume*. This must mean that in going from the activated state to the products the molecules give up, not only the energy absorbed on activation from A to B, but also the energy ΔE corresponding to the difference in the energy levels of C and A. Hence it may be concluded that the concept of an activation energy in no way violates the thermodynamic relations between the energies of products and reactants. All it does is to introduce two thermal quantities such that their difference always leads to the heat of reaction. At the same time it is evident that no information about energies of activation can be obtained from heats of reaction without the aid of kinetic study or other considerations.

The Collision Theory of Bimolecular Reactions. The collision theory of reaction velocity attempts to account for the observed kinetics of reactions in terms of the molecular behavior of the reacting systems. We shall consider this theory first for bimolecular and then for unimolecular reactions. The theory of termolecular reactions is still not sufficiently developed for discussion here.

According to the collision theory of bimolecular reactions two mole-

cules in order to react must (a) collide, and (b) possess on collision sufficient energy for them to be activated. This means that out of all the
collisions occurring only those will be fruitful which involve active
molecules. Consequently, if Z is the number of molecules colliding per
cubic centimeter per second in a reacting system containing 1 mole of
reactant per liter, and q the fraction of these that are activated, then
the specific rate of the bimolecular reaction in molecules per cubic
centimeter per second, k, should be given by

$$k = Zq \qquad (55)$$

For bimolecular *gaseous* reactions Z and q can be calculated with the
aid of the kinetic theory of gases. We have seen in Chapter I, equation
(39), that for a gas containing only *one kind* of molecule the number of
colliding molecules per cubic centimeter per second is

$$Z_1 = \sqrt{2}\, \pi\sigma^2 v n^2 \qquad (56)$$

where σ is the molecular diameter, v the average molecular velocity in
centimeters per second, and n the number of molecules per cubic centimeter. Substituting $v = 0.921\sqrt{(3\,RT/M)}$, from equation (37) of the
same chapter, and collecting constants, the expression for Z_1 becomes

$$\begin{aligned}
Z_1 &= 0.921\sqrt{2}\, \pi\sigma^2 n^2 \sqrt{\dfrac{3\,RT}{M}} \\
&= 6.5 \times 10^4\, \sigma^2 n^2 \sqrt{\dfrac{T}{M}}
\end{aligned} \qquad (57)$$

However, when two *different* molecules are involved, $Z = Z_{1,2}$, and is
given by

$$Z_{1,2} = n_1 n_2 \left(\frac{\sigma_1 + \sigma_2}{2}\right)^2 \sqrt{\frac{8\,\pi RT(M_2 + M_1)}{M_1 M_2}} \qquad (58a)$$

$$= 1.14 \times 10^4 (\sigma_1 + \sigma_2)^2 n_1 n_2 \sqrt{\frac{T(M_2 + M_1)}{M_1 M_2}} \qquad (58b)$$

Here σ_1 and σ_2 are the molecular diameters of the respective molecules,
M_1 and M_2 their molecular weights, n_1 and n_2 the number of molecules
of each type per cubic centimeter at temperature T. $Z = Z_1$ is used
when all the reacting molecules are the same, as in $2\,HI \longrightarrow H_2 + I_2$;
on the other hand, $Z = Z_{1,2}$ is utilized when the two molecules are
unlike, as in the reaction $H_2 + I_2 \longrightarrow 2\,HI$.

The fraction of activated molecules q can be ascertained from the
kinetic theory of gases. The molecules in a gas are in constant motion
during which they enter into repeated collisions with each other. These
collisions cause a redistribution of translational energy, resulting in a

small fraction of the molecules acquiring an energy considerably above the average translational energy of all the molecules. From the Maxwell distribution law for molecular velocities (Chapter I, p. 24) it can be shown that in any gas containing n molecules per cubic centimeter at a temperature T, the number of these, n', which will possess an energy E or higher above the average energy of all the molecules is

$$n' = ne^{-\frac{E}{RT}} \tag{59}$$

If E is identified now with ΔE^*, the energy of activation, the *fraction* of molecules having energies of ΔE^* or higher, i.e., activated, follows from equation (59) as

$$\frac{n'}{n} = e^{-\frac{\Delta E^*}{RT}} = \frac{\text{number of activated molecules}}{\text{total number of molecules}} \tag{60}$$

and this is also q. The expression for k in equation (55) becomes, therefore,

$$k = Ze^{-\frac{\Delta E^*}{RT}} \tag{61}$$

where Z is given by equation (57) for like molecules or by equation (58) for unlike molecules.

We are in a position now to test this theory with an actual calculation, and to compare the result with experiment. For this purpose we shall take the reaction $2\,HI \longrightarrow H_2 + I_2$ at $556°\,K$, for which the observed value of k is 3.5×10^{-7} sec^{-1} (moles/liter)$^{-1}$, and $\Delta E^* = 44{,}000$ calories. For hydrogen iodide, $\sigma = 3.5 \times 10^{-8}$ cm, $M = 127.9$, and since 1 mole of gas per liter contains 6.0×10^{23} molecules per liter, the number per cubic centimeter is $n = 6.00 \times 10^{23}/1000 = 6.0 \times 10^{20}$. Inserting these values of σ, n, T, and M into equation (57), Z_1 follows as

$$Z_1 = 6.51 \times 10^4 (3.5 \times 10^{-8})^2 (6.0 \times 10^{20})^2 \sqrt{\frac{556}{127.9}}$$

$$= 6.0 \times 10^{31} \text{ molecules}$$

Again,

$$q = e^{-\frac{\Delta E^*}{RT}}$$
$$= e^{-\frac{44{,}000}{1.99 \times 556}}$$
$$= 5.2 \times 10^{-18}$$

Therefore,

$$k = Zq = 6.0 \times 10^{31} \times 5.2 \times 10^{-18}$$
$$= 3.1 \times 10^{14} \text{ molecules reacting/cubic centimeter/second}$$

Multiplying this number by 1000 and dividing by Avogadro's number to obtain the number of moles reacting per liter, we find

$$k = \frac{3.1 \times 10^{14} \times 1000}{6.0 \times 10^{23}} = 5.2 \times 10^{-7} \text{ moles/liter/second}$$

as against the experimentally observed value of $k = 3.5 \times 10^{-7}$. Keeping in mind the absolute nature of the calculation and the uncertainties involved, the agreement between theory and experiment must be considered very good.

Similarly satisfactory concordance between the collision theory and experiment has been obtained with several other bimolecular gas reactions, and also with a number of bimolecular reactions in solution by extending to the latter the ideas outlined above.[1] However, there are many reactions, both in the gas phase and in solution,[2] for which the theory yields results that are too high by factors ranging up to 10^9. For such reactions it is customary to write equation (61) in the form

$$k = PZe^{-\frac{\Delta E^*}{RT}} \tag{62}$$

where P, called the *probability factor*, is inserted to allow for the disparity between calculated and observed rate constants. P may have values ranging from unity for reactions obeying the collision theory to about 10^{-9}. It has been suggested that one reason for "slow" reactions, i.e., those with small values of P, is that not only collision between activated molecules may be necessary for reaction, but also a preferred critical orientation of the reacting molecules. If this is the case, not all collisions between activated molecules will result in reaction, and thus the calculated rate will be higher than the observed. On this basis it may be possible to account for probability factors as low as $10^{-3} - 10^{-4}$, but it is very questionable whether such an explanation can account for P's lower than these. In fact, it will be shown that the collision theory omits from consideration an *entropy of activation* which cannot be disregarded.

The Collision Theory of Unimolecular Reactions. Offhand it is difficult to see how the collision theory could possibly be employed to explain the mechanism of unimolecular reactions. In unimolecular processes only one molecule is participating in the reaction, and consequently the question arises: How do molecules in unimolecular reactions attain their energy of activation? The answer to this question now generally accepted was first suggested by Lindemann[3] in 1922. Lindemann

[1] Moelwyn-Hughes, "Kinetics of Reactions in Solution," Oxford University Press, Oxford, 1933, p. 79.
[2] *Ibid.*, p. 111.
[3] Lindemann, Trans. Faraday Soc., **17**, 598 (1922).

pointed out that the behavior of unimolecular reactions can be explained on the basis of bimolecular collisions provided we postulate that a *time lag exists between activation and reaction* during which activated molecules may either react, or be deactivated to ordinary molecules. Worded differently, the Lindemann hypothesis states that in any unimolecular reaction an equilibrium exists between activated and nonactivated molecules; and, out of this equilibrium activated molecules are removed by reaction. The process may be indicated by the equation

$$A + A \rightleftharpoons A + A^* \searrow_{\text{Products}} \tag{63}$$

where A represents inactive and A^* activated molecules. Since the rate will have to be proportional to the concentration of active molecules, we may write for it

$$\frac{dx}{dt} = k_1' C_{A*}$$

Again, for the equilibrium between active and inactive molecules we have

$$K = \frac{C_A C_{A*}}{C_A^2} = \frac{C_{A*}}{C_A}$$

and therefore $C_{A*} = K C_A$. Substituting this value of C_{A*} into the rate expression, we obtain

$$\frac{dx}{dt} = k_1' C_{A*}$$
$$= k_1' K C_A$$
$$= k_1 C_A \tag{64}$$

which is the equation for a unimolecular reaction.

The time lag postulated for unimolecular reactions becomes plausible when the complexity of the gaseous molecules undergoing unimolecular and bimolecular reaction is compared. Many of the molecules reacting bimolecularly are found to contain several atoms only, and to be thus structurally simple. On the other hand, most of the molecules decomposing unimolecularly contain many atoms. In simple molecules it may be expected that any energy acquired on activation will flow rapidly into the bond to be ruptured and cause immediate reaction. However, in a structurally complex molecule some time may be required for the energy to become localized in the bond where fission occurs, occasioning thereby a time lag between activation and dissociation.

When equation (61) is applied to estimate for unimolecular reactions the number of molecules becoming activated due to bimolecular collisions, it is found that the number of molecules calculated is considerably

smaller than the number actually entering reaction. To allow for this fact the suggestion has been advanced [1] that in unimolecular reactions not only are translational energies involved in activation, but also rotational and vibrational. With the latter included, the fraction of the molecules becoming activated is much larger, and the number of molecules activated can be reconciled with the number observed to react in many unimolecular reactions both in the gas phase and in solution.

According to Kassel,[2] the energy of activation of a unimolecular reaction may be defined as the minimum energy per mole, translational, vibrational, and rotational, which a molecule must possess in order to react. Furthermore, this energy must be localized in the weakest bond in the molecule. When such a bond receives energy equal to or in excess of the activation energy, it is loosened or broken, and decomposition results.

One of the interesting consequences of the Lindemann hypothesis is that at low pressures gaseous reactions proceeding by a unimolecular mechanism should pass over to the one for bimolecular processes. The reason for the transformation of mechanism is that at low pressures the time between collisions may become long enough to be comparable to, or greater than, the time lag between activation and reaction. Under these circumstances the rate of activation will be the governing factor, and the rate should follow the bimolecular law. Such changes in molecularity have in fact been observed in a number of unimolecular reactions.

+The Theory of Absolute Reaction Rates. The *theory of absolute reaction rates*,[3] based on statistical mechanics, is an alternate approach to the problem of reaction kinetics. This theory postulates that molecules before undergoing reaction must form an *activated complex* in equilibrium with the reactants, and that the rate of any reaction is controlled by the concentration of the complex present at any instant. The state of affairs in any reaction $A + B$ can be represented by the scheme

$$A + B \rightleftharpoons [A \cdot B]^* \xrightarrow{k} \text{Products} \tag{65}$$

$$\underset{\text{Reactants}}{} \quad \underset{\substack{\text{Activated} \\ \text{complex}}}{}$$
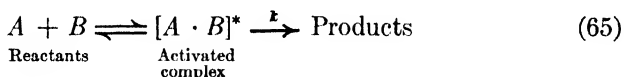
The activated complex is assumed to be endowed with certain properties of an ordinary molecule and to possess some, although temporary, stability. On the basis of these ideas Eyring was able to show that the rate constant k of *any* reaction, no matter what the molecularity or order may be, should be given by

$$k = \frac{RT}{Nh} e^{-\frac{\Delta F^*}{RT}} \tag{66}$$

[1] See Hinshelwood, "Kinetics of Chemical Change in Gaseous Systems," third edition, Oxford University Press, Oxford, 1933.

[2] Kassel, J. Phys. Chem., **32**, 225 (1928).

[3] Eyring, J. Chem. Physics, **3**, 107 (1935); Wynne-Jones and Eyring, *ibid.*, **3**, 492 (1935).

In this equation R is the gas constant in ergs per mole per degree, N Avogadro's number, h is a universal constant, called Planck's constant, equal to 6.62×10^{-27} erg-seconds, T the absolute temperature, and ΔF^* is the *free energy of activation*, i.e., the difference in free energy between the activated complex and the reactants. Since the activated complex is in equilibrium with the reactants, we may resort to thermodynamics and write for ΔF^*,

$$\Delta F^* = \Delta H^* - T\Delta S^* \tag{67}$$

where ΔH^* is the *heat of activation*, which for our purposes may be taken as identical with ΔE^*, and ΔS^* is the *entropy of activation*. Inserting equation (67) into equation (66) we obtain for k

$$k = \frac{RT}{Nh} e^{\frac{\Delta S^*}{R}} c^{-\frac{\Delta E^*}{RT}} \tag{68}$$

Equation (68) is the fundamental relation of this theory. Comparison of this equation with equation (62) of the collision theory shows that

$$PZ = \frac{RT}{Nh} e^{\frac{\Delta S^*}{R}} \tag{69}$$

and consequently the difficulty occasioned in the latter theory by the appearance of the probability factor P can be reconciled by the introduction of the entropy of activation. Although at present it is not possible to estimate theoretically ΔS^* for any but the simplest reactions, it is believed that the theory of absolute reaction rates does represent a step forward in elucidating chemical kinetics. Further, its use has proved fruitful in a number of other directions as well.[1]

Primary Salt Effect in Ionic Reactions. The velocity constants of reactions in solution involving either nonelectrolytes or nonelectrolytes and ions are found to be essentially unaffected by presence of electrolytes in solution. On the other hand, the velocity constants of reactions between ions are sensitive to variation of the ionic strength of the solution, and change with the latter in a manner dependent on the charges of the reacting ions.

To explain the existence of this *primary salt effect*, Brönsted[2] in 1922 suggested a theory of ionic reactions which is in essence a special case of the theory of absolute reaction rates. We shall follow here a treatment due to Bjerrum because it is simpler than Brönsted's. Bjerrum, like Eyring after him, postulates that (*a*) reacting ions go to form an activated complex in equilibrium with the reactants, and (*b*) that the rate

---

[1] See Glasstone, Laidler, and Eyring, "The Theory of Rate Processes," McGraw-Hill Book Company, Inc., New York, 1941.

[2] See Brönsted, Chem. Rev., **5**, 231 (1928), for a very exhaustive treatment of this subject.

of reaction is proportional to the *concentration* of the complex. Thus, any ionic reaction between A^{z_A} and B^{z_B} can be represented analogously to equation (65) by the scheme

$$\underset{\text{Reactants}}{A^{z_A} + B^{z_B}} \rightleftharpoons \underset{\text{Activated complex}}{[(A \cdot B)^{(z_A+z_B)}]^*} \xrightarrow{k'} \text{Products} \qquad (70)$$
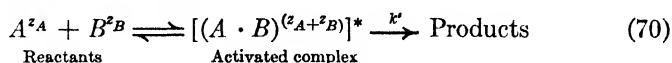
where z_A and z_B are the charges of the two ions, while $(z_A + z_B)$ is the charge on the activated complex. If we let C_X be the concentration of the complex at any instant, then in view of the second postulate the rate at any time t is given by

$$\frac{dx}{dt} = k'C_X \qquad (71)$$

But, since the complex is in equilibrium with the reactants, the thermodynamic equilibrium constant K must be

$$\begin{aligned} K &= \frac{a_X}{a_A a_B} \\ &= \frac{C_X f_X}{(C_A f_A)(C_B f_B)} \end{aligned} \qquad (72)$$

where the a's, C's, and f's represent, respectively, the activities, concentrations, and activity coefficients of the respective species. From equation (72) C_X follows as

$$C_X = (KC_A C_B)\frac{f_A f_B}{f_X}$$

which on substitution into equation (71) yields, for the rate,

$$\begin{aligned} \frac{dx}{dt} &= k'KC_A C_B \cdot \frac{f_A f_B}{f_X} \\ &= \left(k_0 \frac{f_A f_B}{f_X}\right) C_A C_B \end{aligned} \qquad (73)$$

on setting $k_0 = k'K$.

Equation (73) indicates that the experimentally evaluated constant k, as obtained from the ordinary rate equation for the process depicted in equation (70), namely,

$$\frac{dx}{dt} = kC_A C_B$$

is not the true reaction constant k_0, but involves also the activity coefficients of the reactants and the complex. In fact,

$$k = k_0 \frac{f_A f_B}{f_X} \qquad (74)$$

Since the activity coefficients of the charged reactants and complex depend on the ionic strength of the solution, it is to be anticipated that the velocity constant k will also be a function of the ionic strength. By applying the Debye-Hückel theory of activity coefficients, it can be shown that for *dilute* solutions k should vary with ionic strength μ according to the equation

$$\log_{10} \frac{k}{k_0} = 2\, A z_A z_B \sqrt{\mu} \quad (75)$$

where A is the Debye-Hückel constant [1] dependent only on the temperature. According to equation (75), a plot of $\log_{10} k$ or of $\log_{10} k/k_0$ against $\sqrt{\mu}$
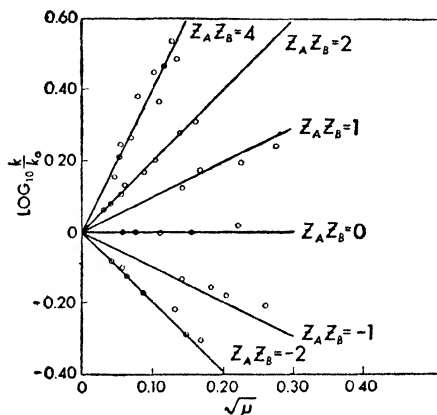
Fig. 7. Primary Salt Effect in Ionic Reactions

should be linear with a slope equal to $(2\, A z_A z_B)$, i.e., *dependent only on the product $z_A z_B$ of the charges of the reacting ions.* That this requirement is in line with observation is indicated by Fig. 7, where a plot of $\log_{10} k/k_0$ vs. $\sqrt{\mu}$ is given for reactions of various $z_A z_B$ values. The solid lines are those predicted by equation (75), while the points are experimental results.

It should be noticed that when $z_A z_B = 0$, as for a reaction between a nonelectrolyte and an ion, the primary salt effect is essentially zero. This will be true for a reaction of the type

$$CH_2ICOOH + CNS^- \longrightarrow CH_2(CNS)COOH + I^-$$
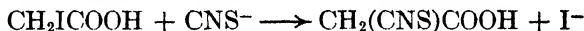
Again, when $z_A z_B$ is positive, as for the reaction

$$CH_2BrCOO^- + S_2O_3^{--} \longrightarrow CH_2(S_2O_3)COO^- + Br^- \quad (z_A z_B = +2)$$
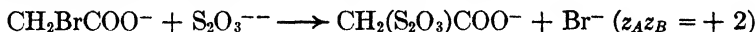
the salt effect will be positive, and k will increase with ionic strength. Lastly, when $z_A z_B$ is negative, as for

$$[Co(NH_3)_5Br]^{++} + OH^- \longrightarrow [Co(NH_3)_5OH]^{++} + Br^- \quad (z_A z_B = -2)$$
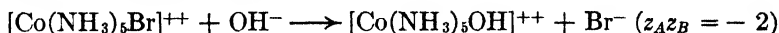
the salt effect is negative, and k will fall with increasing ionic strength.

Catalysis. Except where noted, the reactions considered thus far proceed without any accelerating or decelerating influences under the conditions specified. Yet, it is a well-established fact that the velocity of many a reaction can be changed, and that many a reaction which appears not to go may be effected by introducing certain substances

[1] See equation (29), Chapter XVI.

other than the reactants appearing in the stoichiometric equation of the process. Thus, a pinch of manganese dioxide promotes the evolution of oxygen from potassium chlorate, unsaturated hydrocarbons can be hydrogenated in presence of nickel, sulfur dioxide can be oxidized to sulfur trioxide in the presence of platinum, and iodine accelerates the decomposition of nitrous oxide. In all these processes the foreign substances added remain unchanged at the end of the reaction, and may be used over and over again. *Any substance which influences the rate of a chemical reaction but itself remains unchanged chemically at the end is called a catalyst.* And, the phenomenon of rate acceleration, both negative and positive, is designated as *catalysis.*

The power of a suitable catalyst lies not in its ability to modify the free energies of the reactants or products, but rather in its ability to change the rate at which a reaction determined by the free energy relations can take place. For any reaction to be thermodynamically feasible the change in free energy must be negative. Nevertheless, even with a negative ΔF the rate of transformation may be so slow as to make the reactants appear to be inert. In such cases the purpose of a catalyst is to speed the reaction, and to permit a more rapid approach to equilibrium. From a thermodynamic standpoint, therefore, catalysis introduces no complications into the energy relations of the system. Free energy, heat of reaction, and entropy calculations for a reaction remain the same whether the process takes place per se or is "coaxed along" by a catalyst.

Since a catalyst cannot change the ΔF^0 of a reaction, neither can it modify the equilibrium constant. Consequently if the forward rate is affected, so must also be the reverse rate to the extent that the relation $K = k_1/k_2$ remains valid. In other words, a catalyst must accelerate equally both the forward and reverse reactions. This is found to be the case. Occasionally a catalyst may appear to shift the point of equilibrium, but in such instances careful study will reveal that either the catalyst participates actively as a reactant or some other complications such as salt effects exist which have not been allowed for.

The catalytic activity of a catalyst increases generally with its concentration, although this is not invariably true. Further, the concentration of the catalyst appears in the rate equation. This would suggest that the catalyst participates as a reactant, but that it is regenerated on completion of the sequence of reaction steps. However, this function of a catalyst as a modifier of the mechanism of a reaction is not its only effect. It has also been established that use of a catalyst results usually in a decrease of activation energy; and, as this lowering of the energy barrier between reactants and products makes for higher rate of reaction, positive acceleration of the rate is understandable.

Catalysis may be homogeneous or heterogeneous, depending on whether the catalyst forms a single phase with the reactants or constitutes a separate phase. Here we shall consider homogeneous catalysis in gases and in solution, while heterogeneous catalysis will be discussed in the next chapter.

Homogeneous Catalysis in Gases. Possibly the most familiar example of catalysis of a gas reaction by another gas is the oxidation of sulfur dioxide to sulfur trioxide in presence of nitric oxide in the chamber process for preparation of sulfuric acid. Nitric oxide catalyzes also the oxidation of carbon monoxide to carbon dioxide and the decomposition of nitrous oxide to nitrogen and oxygen. The latter reaction is catalyzed as well by iodine, chlorine, and bromine vapors. Iodine vapors have also been found to be very effective in accelerating the decomposition of various aldehydes and ethers, such as acetaldehyde, methyl ethyl ether, di-ethyl ether, and di-isopropyl ether. Without the catalyst the decomposition of the ethers proceeds unimolecularly with energies of activation of 54,500, 53,000, and 61,000 calories for the ethers in the order mentioned. However, in presence of iodine vapors these decompositions take place by a bimolecular mechanism involving a molecule of the ether and of iodine, and with energies of activation of 38,000, 34,300, and 28,500 calories respectively.

For more extended discussion we may choose as typical of such catalyses the decomposition of acetaldehyde. This decomposition, we have seen, proceeds of itself as a second order reaction with an energy of activation of 45,500 calories. But, when iodine vapors are introduced into the system, it is found that the rate is given by $dx/dt = kC_{\text{acetaldehyde}}C_{\text{I}_2}$, where the constant k is about 10,000 times *larger* at 518° C than the constant for the same reaction in absence of iodine. Further, the energy of activation is reduced to 32,500 calories. To account for these observations it has been suggested that instead of decomposing directly according to

$$2 \text{ CH}_3\text{CHO} \longrightarrow 2 \text{ CH}_4 + 2 \text{ CO}$$
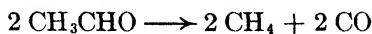
the reaction in presence of iodine vapor takes place in two steps as follows:

$$\text{CH}_3\text{CHO} + \text{I}_2 \longrightarrow \text{CH}_3\text{I} + \text{HI} + \text{CO}$$
$$\text{CH}_3\text{I} + \text{HI} \longrightarrow \text{CH}_4 + \text{I}_2$$
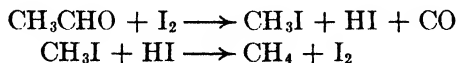
The first of these is considered to be the slower and therefore rate determining. As this alternate mechanism involves a lower energy of activation, the reaction can occur at a faster rate.

Homogeneous Catalysis in Solution. Homogeneously catalyzed reactions in solution are very common, as may be judged from the list given in Table 9. Of particular interest in this connection are the reac-

tions catalyzed by hydrogen or hydroxyl ions, and by acids and bases in the generalized sense. Reactions catalyzed by hydrogen ions only are said to be subject to *specific hydrogen ion catalysis*. For such reactions the rate is proportional to the concentration of the substrate, i.e., the reacting molecule or ion, and to the concentration of the catalyzing hydrogen ions. Again, in *specific hydroxyl ion catalysis* the rate is proportional to the concentrations of substrate and hydroxyl ions. An instance of the latter type of catalysis is illustrated in the data quoted in Table 6 for the hydroxyl ion catalyzed decomposition of diacetone alcohol in water.

TABLE 9

HOMOGENEOUSLY CATALYZED REACTIONS IN SOLUTION

Reaction	Catalyst	Solvent
Decomposition of H_2O_2	HBr, HCl, HI	H_2O
Oxidation of persulfates	Ag^+	H_2O
$Ce^{++++} + S_2O_3^{--}$	I^-	H_2O
Inversion of menthone	$C_2H_5O^-$	C_2H_5OH
Inversion of sucrose	H^+	H_2O
Hydrolysis of esters, amides, acetals	H^+	H_2O
Decomposition of diazoacetic acid	H^+	H_2O
Hydrolysis of sulfamic acid	H^+	H_2O
Hydrolysis of pyrophosphates	H^+	H_2O
Decomposition of nitroso-triacetone-amine	OH^-	H_2O
Decomposition of triacetone alcohol	OH^-	H_2O
Conversion of acetone to triacetone alcohol	OH^-	H_2O
Hydrolysis of ethyl orthocarbonate, ortho-acetate and orthopropionate	Generalized acid	H_2O
Oxidation of phosphorous and hypophosphorous acids by iodine	Generalized acid	H_2O
Rearrangement of N-bromacetanilide	Generalized acid	C_6H_6, C_6H_5Cl
Decomposition of nitramide	Generalized base	H_2O, isoamyl alcohol, m-cresol
Isomerization of nitromethane	Generalized base	H_2O
Mutarotation of glucose	Generalized acid-base	H_2O
Enolization of acetone	Generalized acid-base	H_2O
Acetylation of β-naphthol	Generalized acid-base	CH_3COOH

In distinction to specific hydrogen ion and hydroxyl ion catalysis, we have *generalized acid or generalized base catalysis*. In generalized acid catalysis every specie present in solution which is an acid in the Brönsted definition of the word acts as a catalyst for the reaction. Thus, if a substrate S, subject to generalized acid catalysis, is allowed to react in a solution of acetic acid and sodium acetate in water, we have in this

solution the acids H_2O, H_3O^+, and $CH_3COOH(HA)$, and hence in such a solution the rate of reaction of the substrate will be given by

$$\frac{dx}{dt} = k_{H_2O}C_{H_2O}C_S + k_{H_3O^+}C_{H_3O^+}C_S + k_{HA}C_{HA}C_S$$

$$= C_S(k_{H_2O}C_{H_2O} + k_{H_3O^+}C_{H_3O^+} + k_{HA}C_{HA}) \tag{76}$$

From a knowledge of the sundry concentrations, and by choosing suitable experimental conditions, the various k's, called *catalytic coefficients*, of the respective acids, can be evaluated, and the overall reaction rate can be ascertained. On the other hand, in generalized basic catalysis the substrate is catalyzed by bases in the generalized sense. Thus, if a substrate S, subject to this type of catalysis, is permitted to react in the acetate-acetic acid mixture mentioned above, the bases present are H_2O, OH^-, and acetate ions (A^-), and the rate equation becomes this time

$$\frac{dx}{dt} = k_{H_2O}C_{H_2O}C_S + k_{OH^-}C_{OH^-}C_S + k_{A^-}C_{A^-}C_S$$

$$= C_S(k_{H_2O}C_{H_2O} + k_{OH^-}C_{OH^-} + k_{A^-}C_{A^-}) \tag{77}$$

Again, by utilizing suitable methods the catalytic coefficients of various bases may be found, and from these the total rate of the reaction.

Finally, there are some reactions, such as the enolization of acetone or the mutarotation of glucose, which are subject to both generalized acid and generalized base catalysis, namely, *generalized acid-base catalysis*. In such reactions the overall rate is the sum of the products of the catalytic coefficient, the concentration of substrate, and the concentration of all acid or basic species present; i.e., the total rate equation is the sum of equations of the type of equations (76) and (77).

In generalized acid, base, or acid-base catalysis the catalytic coefficients of various acids and bases vary greatly in magnitude. For generalized acid catalysis the coefficients increase as the acid strength of the catalyzing acid goes up, while in basic catalysis the same parallelism is found with the basic strength of the catalyzing base.

√**Chain Reactions.** In 1906 Bodenstein and Lind, in studying the kinetics of combination of hydrogen and bromine to form hydrogen bromide, found that the rate of the process could not be represented by any equation deduced on a simple basis, but had to be expressed by the relation

$$\frac{dC_{HBr}}{dt} = \frac{k_1 C_{H_2}\sqrt{C_{Br_2}}}{k_2 + \dfrac{C_{HBr}}{C_{Br_2}}} \tag{78}$$

A satisfactory explanation of equation (78) was finally advanced through the efforts of Christiansen, Herzfeld, and Polanyi in 1919–1920 on the

basis of the following mechanism. These authors postulated that the first stage in the reaction is a dissociation of bromine molecules to bromine atoms,

$$Br_2 \longrightarrow 2\,Br \qquad\qquad (a)$$

which was followed next by the slow reaction,

$$Br + H_2 \longrightarrow HBr + H \qquad\qquad (b)$$

and then by the rapid changes,

$$H + Br_2 \longrightarrow HBr + Br \qquad\qquad (c)$$
$$H + HBr \longrightarrow H_2 + Br \qquad\qquad (d)$$
$$Br + Br \longrightarrow Br_2 \qquad\qquad (e)$$
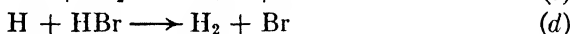
With this mechanism they were able to show that a rate equation of the form of (78) can be deduced, and that consequently the combination of bromine and hydrogen proceeds through a series of consecutive reactions initiated by the appearance of bromine atoms in the system.

A reaction such as the one described proceeding in a series of successive stages initiated by a suitable primary process is called a *chain reaction*. In the hydrogen-bromine and also in the hydrogen-chlorine reaction, which is similar in behavior, the initial step is the appearance of bromine or chlorine atoms. These atoms may result from spontaneous thermal dissociation of the molecules, they may result from decomposition induced by light, or, as Polanyi showed, by the introduction of metallic vapors, such as sodium, which yield atoms through a reaction of the type

$$Na + Cl_2 \longrightarrow NaCl + Cl$$
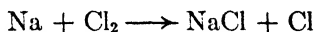
and thus start a chain. In certain other chain reactions the chain may be initiated by the appearance of *free radicals*, such as CH_3, C_2H_5, or CH_3CO, which then react with molecules, but are eventually regenerated to propagate the chain. Finally, a chain may start at the wall of the reacting vessel. This appears to be the case in the reaction of $H_2 + O_2$, $CS_2 + O_2$, and $CO + O_2$, which have been shown to be chain reactions.

Chains initiated in the manner indicated and propagated by atoms or free radicals are called *material chains*. An alternate method of initiation of chain reactions is through "hot molecules." Suppose an atom A reacts with another atom B to yield AB with a large evolution of heat. This liberated energy cannot be removed except by collision with another atom or molecule, which then acquires sufficient energy to become activated. Through a continuation of this process a chain may be started and continued. Chains initiated and propagated through transfer of excess energy are called *energy chains*.

Inspection of equations (a) to (e) for the hydrogen-bromine reaction will reveal that not all the stages of a chain reaction operate to continue a chain once started, but that some of them lead to a termination or "breaking" of chains. Thus, reactions (a) to (d) favor the propagation of the chain by formation of hydrogen or bromine atoms. On the other hand, reaction (e) results in removal of bromine atoms, and hence in chain stoppage. The same would be true of combination of two hydrogen atoms to form a hydrogen molecule. Another very important factor in terminating chains is collision of chain propagators with the walls of the reacting vessel. These collisions may result in either deactivation or reaction of the active agent with the walls. The consequence in either case is a broken chain.

Chain reactions are not limited to the few examples cited. They are encountered in the oxidations of various gaseous hydrocarbons (benzene, ethylene, methane, acetylene), phosphine, and methyl alcohol; in the decomposition of ozone; and in photochemical reactions. Again, indications are that some reactions in solution, namely, oxidation of sodium sulfite by oxygen and the decomposition of certain acids by sulfuric acid, also proceed by a chain mechanism.

Negative Catalysis. The rates of some reactions are readily inhibited by the presence of small quantities of various substances. As examples of this *negative catalysis* may be mentioned the inhibiting action of bromine on the methane-chlorine reaction, and the action of a number of alcohols and aldehydes on the oxidation of sodium sulfite in solution. The readiness with which such substances inhibit the reaction when present in even small quantities suggests that we are dealing here with chain reactions, and that the action of the negative catalysts results from their ability to react with chain propagators to bring about termination of chains. This idea has actually been confirmed by Bäckström for inhibition of the sodium sulfite oxidation, and it is quite probable that a similar explanation may be valid in other cases as well.

It is of interest to point out that the action of "antiknocks" in gasoline, such as tetraethyl lead, is believed to be due to the ability of these substances to decrease the length of chains which develop during oxidation in the engine. Without such inhibitors the chains may reach great length, and cause "knocking" of the engine.

REFERENCES FOR FURTHER READING

1. R. P. Bell, "Acid-Base Catalysis," Oxford University Press, Oxford, 1941.
2. F. Daniels, "Chemical Kinetics," Cornell University Press, Ithaca, N. Y., 1938.
3. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.
4. S. Glasstone, "Recent Advances in Physical Chemistry," The Blakiston Company, Philadelphia, 1936, Chapter V.

5. S. Glasstone, "Recent Advances in General Chemistry," The Blakiston Company, Philadelphia, 1936, Chapters VII and VIII.

6. C. N. Hinshelwood, "Kinetics of Chemical Change in Gaseous Systems," third edition, Oxford University Press, Oxford, 1933.

7. E. A. Moelwyn-Hughes, "Kinetics of Reactions in Solution," Oxford University Press, Oxford, 1933.

8. Schwab, Taylor, and Spence, "Catalysis," D. Van Nostrand Company, Inc., New York, 1937.

PROBLEMS

1. The thermal decomposition of ethylene oxide in the vapor phase may be assumed to proceed as follows:

$$CH_2\text{—}CH_2(g) \longrightarrow CH_4(g) + CO(g)$$
$$\diagdown_O\diagup$$
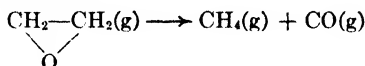
The following data are given by Heppert and Mack [J. Am. Chem. Soc., **51**, 2706 (1929)] at 414.5° C:

Time (min)	0	5	7	9	12	18
Pressure (mm)	116.51	122.56	125.72	128.74	133.23	141.37

Show that the decomposition follows a first order reaction, and calculate the mean specific rate constant.   *Ans.* $k_1 = 0.0123$ min^{-1}

2. From the mean rate constant for the decomposition of azoisopropane at 270° C given in Table 2, calculate (a) the percentage of the original sample decomposed after 25 minutes, and (b) the time required for the reaction to go 95% to completion.

3. An artificially produced radioactive isotope decomposes according to the first order law with the emission of α particles. The half-life period is 15 min. In what time will 80% of the sample be decomposed?

4. At 25° C the half-life period for the decomposition of N_2O_5 is 5.7 hr and is independent of the initial pressure of N_2O_5. Calculate (a) the specific rate constant, and (b) the time required for the reaction to go 90% to completion.

     Ans. (a) 0.122 hr^{-1}; (b) 18.9 hr

5. The first order catalyzed decomposition of H_2O_2 in aqueous solution is followed by titrating the undecomposed H_2O_2 with $KMnO_4$ solution. By plotting the proper function, ascertain from the following data the value of the rate constant:

Time (min)	0	5	10	20	30	50
Ml of $KMnO_4$ for Given Amount of H_2O_2 Solution	46.1	37.1	29.8	19.6	12.3	5.0

     Ans. $k_1 = 0.0437$ min^{-1}

6. Show the steps in the integration of the rate expression:

$$\frac{dx}{dt} = k_2(a - x)(b - x)$$

7. The following table gives kinetic data [Slator, J. Chem. Soc., **85**, 1286 (1904)] for the reaction between $Na_2S_2O_3$ and CH_3I at 25° C, the concentrations being expressed in arbitrary units:

Time (min)	0	4.75	10	20	35	55	∞
$Na_2S_2O_3$	35.35	30.5	27.0	23.2	20.3	18.6	17.1
CH_3I	18.25	13.4	9.9	6.1	3.2	1.5	0

What is the order of the reaction and the mean specific rate constant?

8. The rate of saponification of methyl acetate at 25° C was studied by making up a solution 0.01 molar in both base and ester, and titrating the mixture at various times with standard acid. The following data were thus obtained:

Time (min.)	3	5	7	10	15	21	25
Concentration of Base Found	0.00740	0.00634	0.00550	0.00464	0.00363	0.00288	0.00254

Demonstrate by a graphical method that the reaction is second order, and determine the specific rate constant.

9. From the data of the preceding problem, calculate the time required for the reaction to proceed 95% to completion when the initial concentrations of base and ester are both 0.004 mole/liter. What would the half-life period be in this case? *Ans.* 408 min; 21.5 min

10. The following reaction [Maron and LaMer, J. Am. Chem. Soc., **60**, 2588 (1938)]

$$CH_3CH_2NO_2 + OH^- \longrightarrow H_2O + CH_3CH=NO_2^-$$
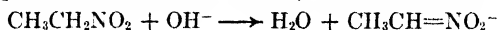
is second order, with the rate proportional to the concentrations of nitroethane and OH^- ions. The specific rate constant at 0° C is 39.1 (moles/liter)$^{-1}$ min^{-1}. After what time will the base be 99% neutralized for a solution containing 0.005 mole/liter of nitroethane and 0.003 mole/liter of NaOH at 0° C?

11. At 25° C the specific rate constant for the hydrolysis of ethyl acetate by NaOH is 6.36 moles^{-1}-min^{-1}. Starting with concentrations of base and ester of 0.02 mole/liter at 25° C, what proportion of ester will be hydrolyzed in 10 min? *Ans.* 56%

12. For the reduction of $FeCl_3$ by $SnCl_2$,

$$2\ FeCl_3 + SnCl_2 \longrightarrow 2\ FeCl_2 + SnCl_4$$
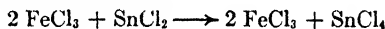
the following data were obtained at 21° C:

t (min)	1	3	7	11	40
x	0.01434	0.02664	0.03612	0.04102	0.05058

where x is the amount of $FeCl_3$ reacted in moles per liter. The initial concentrations of $SnCl_2$ and $FeCl_3$ were respectively 0.0625 and 0.1350 mole/liter. What is the order of the reaction, and the average specific rate constant?

13. The conversion of hydroxyvaleric acid into valerolactone at 25° C in 0.025 N HCl solution was followed by titration with standard base. The data obtained were:

Time (min)	0	48	124	289	∞
cc of Base	19.04	17.60	15.80	13.37	10.71

What is the reaction order and mean specific rate constant?

14. In the following table are listed specific rate constants k for the decomposition of N_2O_5 at various temperatures:

t° C	k (sec^{-1})
0	7.87×10^{-7}
25	3.46×10^{-5}
35	1.35×10^{-4}
45	4.98×10^{-4}
55	1.50×10^{-3}
65	4.87×10^{-3}

Determine graphically the energy of activation, and find the specific rate constant at 50° C.

15. At 378.5° C the half-life period for the first order thermal decomposition of ethylene oxide is 363 min, and the energy of activation of the reaction is 52,000 cal/mole. From these data estimate the time required for ethylene oxide to be 75% decomposed at 450° C. *Ans.* 13.5 min

16. The racemization of pinene, a first order reaction, was studied by D. F. Smith. In the gas phase the specific rate constant was found to be 2.2×10^{-5} min^{-1} at $457.6°$ C and 3.07×10^{-3} at $510.1°$ C. From these data estimate the energy of activation, and the specific rate constant at $480°$ C.

17. The specific rate constants for the second order neutralization of nitroisopropane by a base in aqueous solution are given by the expression [Maron and LaMer, J. Am. Chem. Soc., **60**, 2588 (1938)]:

$$\log_{10} k = \frac{-3163.0}{T} + 11.899$$

The time is in minutes, while the concentrations are in moles per liter. Calculate the energy of activation, and the half-life period at $10°$ C when the initial concentrations of base and acid are each 0.008 mole/liter.

18. The molecular diameters of O_2 and H_2 respectively are 3.39×10^{-8} cm and 2.47×10^{-8} cm. When 1 g of O_2 and 0.1 g of H_2 are mixed in a 1-liter flask at $27°$ C, what will be the number of collisions per second?   *Ans.* 8.76×10^{28}

19. The energy of activation for a bimolecular gaseous decomposition is 20,000 cal/mole. Calculate the fraction of the molecules having sufficient energy to decompose at $27°$ C and at $227°$ C.

20. The gas phase reaction

$$2 A \longrightarrow B + C$$
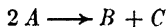
is bimolecular with an activation energy of 24,000 cal/mole. The molecular weight and diameter of A are, respectively, 60 and 3.5 Å. Calculate from kinetic theory the specific rate constant for the decomposition at $27°$ C.

Ans. 2.14×10^{-4} mole^{-1} sec^{-1}

21. From the data given in problem 16 estimate the entropy of activation in the racemization of pinene in calories per mole. Use the average temperature.

Ans. 64.7 eu/mole

22. According to the Brönsted theory, will the specific rate constants of the following reactions remain unchanged, increase, or diminish as the ionic strength is increased:

    (a) $NH_4^+ + CNO^- \longrightarrow CO(NH_2)_2$.

    (b) The saponification of an ester.

    (c) $S_2O_8^{--} + I^- \longrightarrow$ products.

23. The following table gives the specific rate constants (min^{-1}) for the inversion of sucrose in solutions containing various amounts of formic acid, HA, and sodium formate, NaA:

HA	NaA	k
0.25	—	2.55
0.25	0.025	0.72
0.25	0.10	0.24
0.25	0.25	0.118

Assuming the dissociation constant of HA to be 2.1×10^{-4}, and the inversion to be catalyzed by H^+ ions alone, determine whether or not the catalytic coefficient for H^+ remains constant.

Kinetics of Heterogeneous Gas Reactions

Introduction. Although many types of heterogeneous reactions are possible through various combinations of solid, liquid, and gas phases, we shall limit ourselves to a discussion of the kinetics of gaseous reactions which proceed at the surfaces of solid catalysts, i.e., to the *catalysis of gas reactions by solids.* Because such reactions are of great practical importance, they have been investigated more extensively than others, and the principles governing their kinetics are possibly better understood.

Of the importance of heterogeneously catalyzed gas reactions little need be said. Mention of the contact process for manufacture of sulfuric acid, the Haber ammonia process, the oxidation of ammonia to nitric acid, and the catalytic synthesis of methanol from carbon monoxide and hydrogen, among many others, should suffice to make one appreciate the great changes these types of reactions have wrought within the past thirty to forty years in chemical industry in particular, and in national economy in general.

The feature characterizing all gas reactions catalyzed by solids is the fact that the reaction proceeds not in the gas phase but at the surface of the solid catalyst. For this reason this type of catalysis is frequently referred to as *contact catalysis.* In line with present ideas the catalytic reaction is believed [1] to involve at least four distinct steps, namely: (1) adsorption of the reacting gases on the surface of the catalyst; (2) activation of the adsorbed reactants; (3) reaction of the activated gas in the adsorbed phase; and (4) diffusion of the products of the reaction from the surface into the gas phase. Although any one of these processes may be slowest, and therefore rate determining, it has been found possible to approach the kinetics of such reactions from the standpoint of adsorption, and to formulate rate equations in terms of principles based on the Langmuir adsorption isotherm.

Theories of Contact Catalysis. Various theories have been proposed from time to time to explain contact catalysis and to account for observed phenomena. One of these postulates that there is formed be-

---

[1] Schwab, Taylor, and Spence, "Catalysis," D. Van Nostrand Company, Inc., New York, 1937, p. 219.

tween the solid catalyst and one or more of the gaseous reactants an unstable compound, which decomposes rather easily to yield the reaction products and the catalyst. Although this theory finds some application and even substantiation in organic chemistry, as in the Friedel-Crafts reaction, its general application to contact catalysis is attended with so many difficulties that the theory is not favored much at present.

Another theory, due to Bodenstein and Fink, postulates a *thick* layer of reaction products adsorbed on the catalytic surface through which the reactant gases must diffuse to reach the surface and react. Considering the chemical reaction rate at the surface to be rapid, they ascribe the observed rate of reaction to the rate of diffusion of the gases. One of the strongest objections to this theory has been pointed out by Langmuir. Langmuir showed that in order to account for the rate of some reactions it would be necessary to postulate a film of such a thickness as to be visible. As such thick films have never been observed, it is questionable whether this theory is generally applicable in the form proposed.

The theory of contact catalysis which finds general acceptance at present is based on the Langmuir concept of *unimolecular* adsorbed films as developed in Chapter VIII. According to the Langmuir theory a gas is adsorbed on a surface of a solid catalyst in a layer *one molecule deep*. Further, the entire surface need not be covered with adsorbed gas. In fact, parts may be covered and parts may be bare, but it is only the covered portions which will lead to reaction and hence will determine the rate. Consequently, the rate of reaction of a heterogeneously catalyzed gas reaction may be expected to be proportional to the fraction of the surface covered by adsorbed reactants. Concerning the status of this theory Prof. H. S. Taylor [1] of Princeton, in his introductory address to the symposium on "Adsorption of Gases on Solids" held under the auspices of the Faraday Society of England, said: "So far as kinetics is concerned, acceptance of the Langmuir concept may be characterized as universal. As to adsorption, one can summarize the situation by saying that the thick compressed film has during the last decade become progressively thinner, until now the tendency is to reinterpret the idea of the compressed film in terms of the unimolecular layer."

The Langmuir Adsorption Isotherm. It has been pointed out that heterogeneous gas reactions proceed, not in the gas phase, but in the layer adsorbed on the surface of the solid catalyst. If, then, we are to employ the law of mass action to set up the rate equations of such reactions, we must consider as the "active mass" to which the rate is proportional, not the concentration or pressure of the reactants in the gas phase, but their concentration in the adsorbed film. The function

[1] Taylor, Trans. Farad. Soc., **28**, 132 (1932).

of the gas phase is merely to control the concentration of molecules in the adsorbed layer. Since the latter concentration is at any instant proportional to the *fraction of surface covered*, the rate of reaction must also be proportional to this fraction. In other words, we may write in general that

$$\frac{dx}{dt} = k'\theta \tag{1}$$

where θ is the fraction of surface covered by reactant, while k' is a proportionality constant.

To obtain θ as a function of the pressure of the reacting gases, the Langmuir adsorption isotherm, based on the unimolecular layer concept, is utilized. We have seen in Chapter VIII, equation (3), that, for any adsorbed gas, θ, the fraction of surface covered by the gas, is given by

$$\begin{aligned}
\theta &= \frac{k_1 P}{k_2 + k_1 P} \\
&= \frac{bP}{1 + bP}
\end{aligned} \tag{2}$$

where $b = k_1/k_2$ is a constant while P is the pressure of the gas in contact with the adsorbent. This is the expression for θ to be inserted in equation (1). However, in two special limiting cases equation (2) can be simplified as follows. When a gas is only weakly adsorbed b is small, and bP may be disregarded compared to unity. Under these conditions, θ becomes

$$\theta = bP \tag{3}$$

i.e., *the fraction of the surface covered when a gas is only slightly adsorbed is directly proportional to the pressure.* Again, when a gas is strongly adsorbed, as when either b or P or both are large, then $bP \gg 1$, and equation (2) reduces to

$$\theta = \frac{bP}{bP} = 1 \tag{4}$$

Therefore, *when a gas is strongly adsorbed the surface is practically completely covered with a film of gas at all times, and $\theta = 1$.*

Another deduction which can be made from equation (2) is of importance for our purposes. Since θ is the fraction of surface covered, $(1 - \theta)$, the fraction of surface bare, is

$$\begin{aligned}
(1 - \theta) &= 1 - \left(\frac{bP}{1 + bP}\right) \\
&= \frac{1 + bP - bP}{1 + bP} \\
&= \frac{1}{1 + bP}
\end{aligned} \tag{5}$$

For strong adsorption $bP \gg 1$, and hence,

$$(1 - \theta) = \frac{1}{bP}$$

$$= \frac{b'}{P} \tag{6}$$

From equation (6) we see that when a gas is strongly adsorbed on a surface *the fraction* of the latter still *free and available for adsorption of other gases is inversely proportional to the pressure of the adsorbed gas.*

With the aid of these equations we are prepared now to illustrate the kinetics of heterogeneous reactions in a number of simple cases. These illustrations will be confined to reactions in which only a single gas participates as a reactant. For more complex cases involving two gases the student may refer to special books on the subject.[1]

One Reactant Gas Slightly Adsorbed. Consider a gas A decomposing on the surface of a solid catalyst according to

$$A \longrightarrow \text{Products} \tag{7}$$

Let θ_A be the fraction of the surface covered by A at time t and pressure P. Then according to equation (1) the rate of the reaction at any time t should be

$$\frac{dx}{dt} = k'\theta_A$$

When A is only slightly adsorbed, θ_A is given by equation (3). Hence,

$$\frac{dx}{dt} = k'bP = kP$$

$$= k(P_i - x) \tag{8}$$

where $P_i - x = P$, while $k = k'b$ is the *velocity constant* of the reaction. P_i is the initial pressure and x the decrease in it at any time t. On integration equation (8) becomes

$$k = \frac{2.303}{t} \log_{10}\left(\frac{P_i}{P_i - x}\right) \tag{9}$$

which is identical in form with the equation for a first order homogeneous reaction. For this reason equations (8) or (9) are said to represent *first order* heterogeneous reactions.

Quite a number of reactions have been found to follow equation (9). Among these may be mentioned the decomposition of phosphine on glass,

[1] Schwab, Taylor, and Spence, *loc. cit.*
Hinshelwood, "Kinetics of Chemical Change in Gaseous Systems," third edition, Oxford University Press, Oxford, 1933.

porcelain, and silica; the decomposition of formic acid vapors on glass, platinum, and rhodium; the decomposition of nitrous oxide on gold; and the dissociation of hydrogen iodide on platinum. In Table 1 are shown some data for the dissociation of nitrous oxide on gold [1] which may be considered typical of all such reactions.

TABLE 1

DECOMPOSITION OF N_2O ON GOLD AT 900° C

t (sec)	$P_i - x$ (mm)	$k = \dfrac{2.303}{t} \log_{10} \dfrac{P_i}{P_i - x}$
0	200	—
900	167	2.01×10^{-4}
1800	136	2.15
3180	100	2.18
3900	86	2.17
4800	70	2.19
6000	54	2.18
7200	44	2.10

One Reactant Gas Strongly Adsorbed. Since for a strongly adsorbed reactant $\theta \equiv 1$, the rate of the reaction is given by

$$\frac{dx}{dt} = k \qquad\qquad (10)$$

i.e., the rate of the reaction is *constant and independent of pressure.* A reaction obeying equation (10) is said to be of *zero order.* On integration, equation (10) becomes

$$x = kt$$

or

$$k = \frac{x}{t} \qquad\qquad (11)$$

Hinshelwood and Burk [2] found the decomposition of ammonia on tungsten to be practically independent of the initial pressure, and to proceed in accordance with equation (10). This may be seen from the results quoted in Table 2. The same behavior has been observed in the decomposition of ammonia on the surfaces of molybdenum and osmium, and in the dissociation of hydrogen iodide on gold.

One Reactant Gas Moderately Adsorbed. When the reactant is moderately adsorbed, the rate of reaction is proportional to the fraction of surface covered as given by the full equation (2), namely,

$$-\frac{dP}{dt} = k'\left(\frac{bP}{1 + bP}\right) = \frac{kP}{1 + bP} \qquad\qquad (12)$$

[1] Hinshelwood and Pritchard, Proc. Roy. Soc., **108A**, 211 (1925).
[2] Hinshelwood and Burk, J. Chem. Soc., **127**, 1116 (1925).

TABLE 2

DECOMPOSITION OF NH$_3$ ON TUNGSTEN AT 856° C
(P_i = 200 mm)

t (sec)	x (mm Hg)	$k = \dfrac{x}{t}$
100	14	0.14
200	27	0.14
300	38	0.13
400	48	0.12
500	59	0.12
1000	112	0.11

The equation can frequently be approximated by the expression

$$\frac{-\,dP}{dt} = kP^n \tag{13}$$

which makes the rate proportional to the amount adsorbed as given by the Freundlich isotherm. Stock and Bodenstein [1] used equation (13) to represent the course of decomposition of stibine on antimony at 25° C, while equation (12) was found to hold in the decomposition of nitrous oxide on indium oxide.[2]

Retarded Reactions. In some reactions it is observed that not only the reactant but also a product of the reaction is adsorbed by the catalytic surface. In fact, the product may be adsorbed more strongly than the reactant. Since a strongly adsorbed product diminishes the free surface available for condensation of reactant, the fraction of the surface to be covered by the latter must decrease, and thereby also the rate of reaction. In other words, *the effect of a strongly adsorbed product is to retard the reaction producing it.*

Consider specifically the reaction

$$A \longrightarrow B + C \cdots \tag{14}$$

where the reactant A is weakly adsorbed, while the product B is strongly adsorbed. Let P_A and P_B be the pressures of A and B respectively at time t, and let $(1 - \theta_B)$ be the fraction of the surface not covered by B. The rate of transformation of A will then be proportional to the rate at which molecules of A can become attached to the surface, namely, to the fraction of surface exposed and the pressure of A

$$\frac{dx}{dt} = k'(1 - \theta_B)P_A \tag{15}$$

[1] Stock and Bodenstein, Berichte, **40**, 570 (1907).
[2] Schwab, Staeger, and v. Baumbach, Z. physik. Chem., **21B**, 65 (1933).

But when a gas is strongly adsorbed, $(1 - \theta)$ is given by equation (6), which in this case takes the form $(1 - \theta_B) = b'/P_B$. Substituting this value of $(1 - \theta_B)$ into equation (15) we obtain, for the rate of reaction,

$$\frac{dx}{dt} = \frac{k'b'P_A}{P_B}$$
$$= \frac{kP_A}{P_B} \tag{16}$$

i.e., the rate is directly proportional to the pressure of the reactant, and *inversely* to the pressure of the product present which is responsible for the retardation.

The decomposition of ammonia on platinum at 1138° C was found by Hinshelwood and Burk [1] to conform essentially to equation (16). Nitrogen appears to have no effect on the course of the reaction, but the hydrogen formed exhibits a strong retardation, as may be seen from Table 3. In this table column 1 gives the decrease in the pressure of ammonia produced in 120 sec on starting in each case with 100 mm, while column 2 gives the pressures of added hydrogen present initially. For a threefold increase in the pressure of the retarding gas the decomposition is decreased in an equal period from 33 to 10 per cent of the total.

TABLE 3

RETARDATION OF NH_3 DECOMPOSITION ON Pt AT 1138° C BY H_2
$(P_i = 100 \text{ mm})$

P_{NH_3} decomposed in 120 sec (mm)	P_{H_2} initially added (mm)
33	50
27	75
16	100
10	150

When the product responsible for the retardation is moderately rather than strongly adsorbed, equation (15) is still applicable, but $(1 - \theta_B)$ is given by equation (5). For this case we obtain, therefore,

$$\frac{dx}{dt} = \frac{kP_A}{1 + bP_B} \tag{17}$$

This equation has been applied to the decomposition of nitrous oxide on platinum, cadmium oxide, cupric oxide, and nickel oxide. The retarding gas is oxygen.

The Order of Heterogeneous Reactions. The order of a heterogeneous reaction is defined as the total power to which the pressures

[1] Hinshelwood and Burk, J. Chem. Soc., **127**, 1114 (1925).

appear in the rate equation. A more suitable criterion of the order n of a heterogeneous reaction is the period of half-life, which is inversely proportional to the $(n - 1)$ power of the initial pressure. By studying the period of half-life of a reaction as a function of the initial pressure, it is possible to evaluate $n - 1$, and hence n, the reaction order.

The value of n so obtained is not necessarily the *true order* of the reaction as it takes place on the surface. Rather, it is the *apparent order* as ascertained by following the surface reaction indirectly through its influence on the pressure of the gas phase. As we have seen in the examples cited, n may or may not be a whole number. When a single reacting gas is weakly adsorbed, the reaction is first order and obeys an equation identical in form with a homogeneous first order reaction. Again, when a single reactant is strongly adsorbed, $n = 0$, and the reaction proceeds independently of the pressure. On the other hand, for the intermediate adsorption cited for the decomposition of stibine on antimony, $n = 0.6$. The last two results would be hard to understand unless it is remembered that in heterogeneous reactions only apparent orders are obtained, and that changes may proceed on the surface with orders which may be quite different.

For a single reacting gas only slightly adsorbed, and where products exhibit no retardation, the apparent and true orders may be inferred to be identical. For fractional and zero order reactions the true and apparent orders cannot possibly be the same. In such cases the true order must be deduced from the nature of the reaction and from any other data available. This correlation of reaction orders is further complicated by adsorption of and retardation due to products. In general, the effect of a strongly adsorbed product, i.e., to the extent that the free surface is inversely proportional to its pressure, is to yield an apparent order *one less than the true order*. This can readily be shown by integrating the rate equation of such a process and solving for the time of half-life. Consequently, when an apparent zero order reaction is found to be retarded by a strongly adsorbed product, it may be assumed that the reaction has a true first order.

Effect of Temperature on Heterogeneous Reactions. As for homogeneous reactions, the influence of temperature on the rate constant k of heterogeneous reactions is given by the Arrhenius equation,

$$\frac{d \ln k}{dT} = \frac{\Delta E_a^*}{RT^2} \tag{18}$$

However, ΔE_a^* is called here the *apparent energy of activation*. The apparent energy of activation, evaluated from the observed velocity constants, is not necessarily the energy required to activate the reactants on the *surface*, which is the true activation energy. The latter may

be modified by the heats of adsorption of reactants, or reactants and products, to yield an apparent activation energy which may be quite different from the true. For a discussion of the relation between these two activation energies the student must be referred to Hinshelwood.[1]

The energy of activation appears to be intimately connected with the functions of the catalyst in heterogeneous reactions. Comparison of the energies of activation of gaseous reactions at some surface with the energies of activation of the same reactions proceeding homogeneously almost invariably indicates that the energy of activation of the heterogeneous reaction is considerably lower than that of the homogeneous. This may be seen from Hinshelwood's figures quoted in Table 4. These results point to the fact that the action of the catalyst leads to a *lowering of the activation energy*. With a lower energy barrier more molecules may become active and enter into reactions than is possible in reactions with a high energy threshold. The mechanism by which this lowering is accomplished is as yet not well understood. Hinshelwood, reasoning from the above table, is impressed with the fact that, while the homogeneous reactions are bimolecular, the heterogeneous reactions mentioned are probably in every instance unimolecular. Instead, therefore, of requiring for activation, say in the case of hydrogen iodide, 44,000 calories to be shared between two molecules before reaction may occur, on the surface of gold a single molecule may decompose as soon as it has acquired 25,000 calories. As it is more probable for a single molecule to attain an energy level of 25,000 calories than it is for two to share 44,000 calories between them, we should expect in the monomolecular reaction a greater fraction of active molecules on the surface, and hence a more rapid rate of decomposition of hydrogen iodide than in the homogeneous reaction. This appears to be the case. However, whether this explanation is generally applicable still remains to be established.

TABLE 4 [2]

COMPARISON OF ENERGIES OF ACTIVATION OF HOMOGENEOUS AND HETEROGENEOUS REACTIONS

Decomp. of Gas	ΔE^* Homog. (bimolecular)	ΔE_a^* Heterog. (probably unimolecular)
HI	44,000	25,000 (gold)
N_2O	58,500	29,000 (gold)
N_2O	58,500	32,500 (platinum)
NH_3	80,000 (est.)	39,000 (tungsten)

Catalytic Poisons. Small quantities of foreign substances added to a reacting system are frequently sufficient to impair seriously the cata-

[1] Hinshelwood, "Kinetics of Chemical Change in Gaseous Systems," third edition, Oxford University Press, Oxford, 1933.
[2] Hinshelwood, *loc. cit.*

lytic activity of a surface. Such substances are termed *catalytic poisons*. Poisons may be of two kinds, temporary and permanent. In temporary poisoning the surface decreases in activity, or loses it entirely, only for the period that the poison is in contact with it. As soon as the poison is removed from the presence of the catalyst, the activity is restored. The diminution in activity is due in most instances to a strong preferential adsorption of the poison rather than reactants on the catalyst surface. With sufficiently strong adsorption the reactant may be displaced completely, and the entire surface may be covered with an inactive blanket of the poison.

An illustration of temporary poisoning is the retarding effect exercised by carbon monoxide on the hydrogenation of ethylene in presence of copper.[1] Into this class fall also the reactions in which retardation by products takes place. In expressing the rates of such reactions the pressure of the poison appears in the denominator of the rate equation whether the poison takes part in the reaction or not.

Permanent poisoning, on the other hand, involves a chemical interaction between the surface and the poison to form a new surface catalytically inert. Activity can be reestablished only by chemical rejuvenation. Among the permanent poisons volatile silicon and sulfur compounds are particularly fatal to the life of many catalysts. So are arsenic compounds, especially to platinum.

The very small quantities sufficient to poison a catalytic surface have led H. S. Taylor to suggest that only a fraction of the total exposed surface is catalytically active. He, as well as others, showed that the quantity of poison effective in stopping activity is sometimes too small to cover the whole surface with a film one molecule deep. In his theory of "active centers" Taylor accounts for this by considering the surface as irregular, with high catalytic activity localized in certain spots that are more elevated than others. As soon as those spots become covered with poison, the activity of the surface diminishes to practically zero. Furthermore, since those spots are only a small part of the total surface, the amount of substance necessary to poison a catalyst is correspondingly small.

Catalytic poisoning is occasionally very specific and may be taken advantage of to control the products of a reaction. For instance, the decomposition of alcohol in presence of copper proceeds in two stages,

$$CH_3CH_2OH \longrightarrow CH_3CHO + H_2 \qquad (a)$$
$$CH_3CHO \longrightarrow CH_4 + CO \qquad (b)$$

the first of which is uninfluenced by water vapor, while the second is considerably retarded. Therefore, by using alcohol with some water in

---

[1] Pease, J. Am. Chem. Soc., **45**, 1196, 2235 (1923); Pease and Stewart, *ibid.*, **47**, 1235 (1925).

it, Armstrong and Hilditch[1] were able to increase the yield of acetaldehyde and to prevent materially at the same time the formation of undesired methane. This type of selective poisoning is of great industrial importance.

Promoters. Whereas some substances decrease the catalytic activity of a surface, others, known as *promoters*, increase it. The promoter itself usually is noncatalytic and consists in most instances of a metal or highly fusible metallic oxide incorporated into the body of the catalyst. The amount of promoter required is small. In the few quantitative investigations of the effect of promoters, the increase in velocity proved considerably greater than would correspond to any increase in total surface area. Thus, Russell and Taylor found that for a 20 per cent increase in surface due to promoters the rate of reaction increased tenfold. This observation and others point to an actual improvement in the quality of a surface by promoters, rather than to a mere evolution of surface.

No satisfactory and generally acceptable explanation of promoter action has yet been advanced.

Nature of Catalysts. Technical catalysts vary greatly both in their chemical nature and in the form in which they are used. Chemically, catalysts may be pure metals, metallic oxides, or salts. They may further be individual substances, like nickel in hydrogenation processes, or a mixture of several substances, as Fe_2O_3–Bi_2O_3 in the oxidation of ammonia to nitric acid. The mixed catalysts, because of promoter action exerted by small quantities of a second or third substance, are as a rule more active than the individual components. The mixtures need not necessarily be composed of all metals or all oxides, but may consist of metals and oxides, oxides and salts, etc. The effect produced by addition of a substance to a catalyst cannot be foretold. Each and every combination must be prepared and tried under operating conditions.

Catalysts may further be *supported* or *unsupported*. An unsupported catalyst is one used by itself without any mounting. Supported catalysts, in turn, are those which, either because of added strength or increased surface, are mounted on some base. The latter is usually an inert porous substance with large exposed surface upon which the catalyst is deposited. The supports used most frequently are asbestos, kieselguhr, powdered pumice, porous firebrick, silica and other gels, and various salts. A supported catalyst familiar to every student is platinized asbestos.

The activity of catalysts depends also to a large degree on the method and the care exercised in their preparation. Generally only the purest

[1] Armstrong and Hilditch, Proc. Roy. Soc., **97A**, 262 (1920).

chemicals can be employed, the preparative process must yield the catalyst in finely divided form, and the temperature must be carefully controlled to avoid sintering of the fine mass with attendant reduction in surface area.

Some Important Catalysts. Such a wide variety of substances may act catalytically in various reactions that it is difficult to predict *a priori* which may or may not act as catalysts. Yet a number of generalizations have emerged from experience, and these are of great aid in any search for a catalyst.

The synthetic preparation of ammonia is catalyzed by many metals. As catalysts effective to a greater or lesser degree may be mentioned osmium, uranium, iron, nickel, manganese, molybdenum, tungsten, and cobalt. The activity of metals such as molybdenum, tungsten, iron, nickel, and cobalt is further enhanced by the addition of some high-melting and difficultly fusible oxides, especially those of aluminum, magnesium, chromium, manganese, and the rare earths. Thus the catalyst most widely used at present is iron with an admixture of considerable aluminum oxide and some potassium oxide. It is cheap, easily available, and of satisfactory activity.

In various hydrogenation processes, both in the liquid and gaseous phases, the most active catalyst is nickel. Others in order of decreasing activity are cobalt, platinum, palladium, iron, and copper. For vigorous hydrogenation nickel is most commonly used, and occasionally also cobalt. The latter, however, is less active and as a whole less desirable. Iron can be used on account of its more moderate activity to hydrogenate unsaturated organic compounds, aldehydes, and ketones. However, it will not hydrogenate the benzene ring. Finally, copper, with still more moderate activity, is useful as a catalyst for milder reductions, such as that of nitrobenzene to aniline, and in fact most nitro compounds to amines. Another catalyst with properties similar to copper is tin.

Nickel and copper are also suitable catalysts for the dehydrogenation of alcohols in the vapor phase to aldehydes and ketones. As promoters for these catalysts have been used thorium, aluminum, tungsten, and chromium oxides and kaolin ($Al_2O_3 \cdot 2 SiO_2 \cdot 2 H_2O$). The oxides mentioned are also effective by themselves in the removal of water from alcohols to form olefins.

Catalysis in Industry. Although considerable progress has been made toward a better understanding of catalytic phenomena, theory has lagged far behind practice in the development of catalysts for various reactions. Most of the information available at present on technical catalysts, their specificity, preparation, activity, and life, has been obtained by purely empirical means. This vast information has been collected painstakingly by "cut and try" methods, and from it have

been selected the substances which may be referred to as "good catalysts" for various reactions. From a chemical standpoint a good catalyst is any substance which improves the course of a reaction or makes a simpler one possible. From the technical standpoint, however, the catalyst must also meet another condition — its use for the preparation of a product must offer distinct economic advantages in the saving of raw materials, use of cheaper materials, or conservation of energy.

The technical development of catalysts has made tremendous strides within the past 30 years and has led to the improvement of many old and the introduction of many new chemical processes. The progress, however, has been purely empirical. From the great wealth of information collected very few significant theoretical generalizations have been evolved. But it is precisely these generalizations which are badly needed to allow a more economical improvement of known catalysts and discovery of still unknown ones. When a principle is found which could foretell the type of catalyst suitable for a certain type of reaction, technical progress will be much more rapid, and with less waste in time, money, and effort.

REFERENCES FOR FURTHER READING

1. K. C. Bailey, "The Retardation of Chemical Reactions," Edward Arnold & Co., London, 1937.
2. Berkman, Morrell, and Egloff, "Catalysis," Reinhold Publishing Corporation, New York, 1940.
3. R. H. Griffith, "The Mechanism of Contact Catalysis," Oxford University Press, Oxford, 1936.
4. C. N. Hinshelwood, "Kinetics of Chemical Change in Gaseous Systems," third edition, Oxford University Press, Oxford, 1933.
5. V. N. Ipatieff, "Catalytic Reactions at High Temperatures and Pressures," The Macmillan Company, New York, 1936.
6. Marek and Hahn, "Catalytic Oxidation of Organic Compounds in the Vapor Phase," Reinhold Publishing Corporation, New York, 1932.
7. National Research Council, "Twelfth Report of the Committee on Catalysis," John Wiley & Sons, Inc., New York, 1940.
8. Schwab, Taylor, and Spence, "Catalysis," D. Van Nostrand Company, Inc., New York, 1937.

PROBLEMS

1. The diameter of an O_2 molecule is 3.39×10^{-8} cm. Assuming a unimolecular layer, and that the effective area occupied by the molecules is the square of the diameter, how much O_2 could be adsorbed by a W filament whose diameter is 0.02 cm and whose effective length is 10 cm? What pressure would be exerted by this amount of O_2 when contained in a 100-cc flask at 27° C?

$Ans.$ 9.07×10^{-10} moles; 1.69×10^{-4} mm Hg

2. Langmuir [J. Am. Chem. Soc., **40,** 1361 (1918)] measured the volume of various gases adsorbed at 20° C by cover glasses having a total surface area of 1966.0 sq cm with the following results:

	Volume (mm³)	Molecular Diameter (Å)
H_2O	354	2.20
CO_2	64.0	4.18
N_2	49.0	3.75

Assuming that each molecule occupies an area equal to the square of its diameter, calculate in each case the molecular thickness of the surface layer.

3. Calculate the pressure of N_2O in contact with a gold surface at 900° C after 2 1/2 hr if the initial pressure is 350 mm. After what time will the decomposition be 95% complete? Use the data given in Table 1.

Ans. 51 mm; 233 min

4. Show how equation (13) may be integrated in order to obtain k in terms of P, t, n, and the initial pressure P_i.

5. Stock and Bodenstein give the following data for the decomposition of SbH_3 on Sb at 25° C.

Time (min)	0	5	10	15	20	25
$P_{stibine}$	1.000	0.731	0.509	0.327	0.189	0.093

Show that the decomposition may be represented satisfactorily by the rate expression $-dP/dt = kP^n$, where $n = 0.6$, and determine the rate constant.

6. Kunsman [J. Am. Chem. Soc., **50,** 2100 (1928)] reports the following data for the decomposition of NH_3 on W at 1100° K:

Initial Pressure of NH_3 (mm)	265	130	58	16
Half-Life Period (min)	7.6	3.7	1.7	1.0

Show that the reaction is approximately zero order, and calculate the mean specific rate constant.

7. From the data given in Table 2 estimate the percentage of NH_3 decomposed after 1 hr, and the time required for the decomposition of 75% of the original sample.

Ans. 100%; 1154 sec

8. According to Hinshelwood the decomposition of N_2O on Pt follows the rate equation

$$\frac{-d(a-x)}{dt} = \frac{k(a-x)}{1+bx}$$

where a is the initial pressure of N_2O in mm, x is the decrease in pressure of N_2O in time t, while k and b are constants. On integration this rate equation yields

$$k = \frac{1+ab}{t}\ln\frac{a}{a-x} - \frac{bx}{t}$$

From the following data obtained at 741° C with $a = 95$ mm:

t (sec)	315	750	1400	2250	3450
x (mm)	10	20	30	40	50

determine the constants k and b by a graphical method.

Ans. $k = 3.99 \times 10^{-4}$; $b = 0.0295$

9. Derive the rate expression for the kinetics of a heterogeneous reaction in which two reactants are moderately adsorbed, and the products do not retard the reaction.

10. Hinshelwood and Topley [J. Chem. Soc., **125**, 393 (1924)] found that the decomposition of phosphine on fused silica followed the first order law. The following data are given for the half-life periods at various temperatures:

$T°K$	828	894	956	1007
Half-Period (sec)	580	117	22	6.3

Determine graphically the apparent energy of activation and estimate the half-life period at 800° K. $\qquad$ *Ans.* $\Delta E^* = 41,200$ cal; 1470 sec

11. Formic acid decomposes into CO_2 and H_2 on a gold surface, following a first order law. Hinshelwood and Topley [J. Chem. Soc., **123**, 1014 (1923)] observed a rate constant of 5.5×10^{-4} at 140.0° C and 9.2×10^{-3} at 185.0° C. Calculate the apparent energy of activation.

Atomic Structure and Radioactivity

Introduction. The first modern attempt at an explanation of the reactivity of chemical elements is Dalton's atomic theory formulated in 1808. To account for the law of definite proportions in chemical combination Dalton postulated that every element consists of indivisible particles, called atoms, each of which has the same mass for a given element, and that these atoms react with each other to form compounds. In terms of such atoms he was able to interpret the law of definite and multiple proportions, and also Lavoisier's finding that the total masses of the products of a reaction were equal to the total masses of the reactants, i.e., the *law of conservation of mass.*

However, attempts to apply Dalton's theory to Gay-Lussac's law of volumes, namely, that gases combine with each other in simple proportions by volume, met with failure because no distinction was made between atoms and molecules. This misconception was first clarified by Avogadro in 1811, when he enunciated his famous hypothesis that at constant pressure and temperature equal volumes of all gases contain the same number of molecules — and not necessarily atoms. With the aid of this principle Cannizzaro in 1858 was able to show how molecular weights may be estimated, and that the law of combining volumes is explicable when it is realized that molecules of an element may consist of more than one atom. This principle permits also a differentiation between atomic and molecular weights of an element, and removes thereby a great deal of the uncertainty that was involved in early assignments of atomic weights.

Dalton, in setting up the atom as the unit of elementary chemical combination, made no effort to establish any relation between the atoms of various elements. This was first done in 1815 by Prout who, reasoning from the proximity of the atomic weights of several elements to whole numbers on the basis of hydrogen as unity, suggested that all elements were composed of multiples of hydrogen atoms. However, further atomic weight determinations, and particularly the atomic weights of such elements as chlorine, 35.46, showed marked differences in atomic weights from whole numbers, so this principle fell into dis-

repute. Only when the existence of isotopes was established at a much later date, and the true significance of deviations from close proximity to whole numbers was appreciated, did Prout's hypothesis become re-established as an important principle in considerations of atomic constitution.

Another important principle which indicated the existence of a close relationship between the elements was the discovery of the periodicity of the elements independently by Mendeleeff (1869) in Russia and Lothar Meyer (1870) in Germany. This discovery, that the chemical and physical properties of the elements are periodic functions of their atomic weights, tended to point to regularities in the structure of atoms, and to the fact that certain structures tended to repeat themselves periodically to yield similarity of chemical and physical properties. Nevertheless, it was not until much later that the reason for this behavior was found, and it was discovered that the regularity does not follow atomic weight but *atomic number*, i.e., the numerical sequence of an element in the periodic table.

Finally, it remained for the electrical researches of Nicholson and Carlisle, Davy, Berzelius, and Faraday to show that matter and electricity are intimately associated, and that electricity itself is corpuscular in nature. These findings, followed by studies of electric discharges through rarefied gases and the discovery of radioactivity, eventually established that atoms are divisible and that on subdivision all atoms can be shown to consist of the same structural units. At present four of these atomic structural units are known: (1) the electron, (2) the proton, (3) the neutron, and (4) the positron. A description of the identification of these units, their properties, and the manner in which they enter into the makeup of atoms, as far as is known, will be the concern of this chapter.

The Electron. As early as 1874, G. Johnstone Stoney pointed out that in order to account for Faraday's laws of electrolysis it is necessary to suppose that electricity as well as matter is atomic in nature. This opinion was emphasized further by Helmholtz in 1881 for both negative and positive electricity. Still, the generality of this view did not become manifest, and the units of electricity were not defined, until J. J. Thomson and his collaborators at Cambridge in England began their famous experiments on the passage of electricity through rarefied gases.

Gases are as a rule poor conductors of electricity. However, when a tube filled with gas is evacuated to pressures of 0.01 mm or lower, and an electric potential is applied across a pair of electrodes sealed into the tube, a discharge takes place between the electrodes during which a stream of rays, called *cathode rays*, is found to be emitted *from the cathode*. These rays travel in straight lines perpendicular to the cathode surface,

they produce a temperature rise in any object they strike, they pass through thin films of metals interposed in their path but are stopped by thicker foils, and they can make any opaque object placed in their path cast a sharp shadow. When they strike the wall of the tube, a fluorescence is produced. The fact that these rays can cast a sharp shadow suggests that they consist of material particles and not electromagnetic radiation. Again, these particles can be deflected by electric and magnetic fields, indicating that they are electrically charged; and, further, the direction of deflection is always such as to indicate that these particles always bear a *negative charge*. Finally, no matter what the nature of the cathode or the gas in the tube may be, the particles are always the same.

These particles constituting the cathode rays have been named *electrons*. The fact that the nature of electrons is independent of the nature of the source from which they come suggests that they are constituents of all matter. To study more precisely the properties of
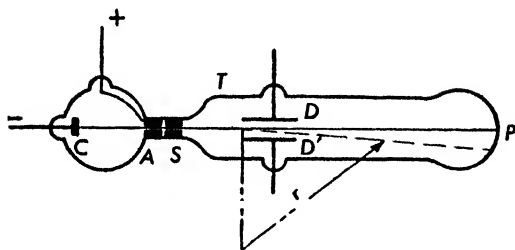
Fig. 1. J. J. Thomson's Apparatus for Determination of *e/m* of Electron

electrons Thomson utilized a combination of electric and magnetic fields placed across the path of these through the discharge tube. His apparatus, shown in Fig. 1, consisted of a tube *T*, into which were sealed the cathode *C*, acting as the source of electrons, the anode *A*, a slit system *S* for defining the electron beam, and a pair of electrodes *D* and *D'* for applying an electric field. The electromagnet for supplying the magnetic field (not shown) was placed outside the tube. A beam of electrons emanating from *C* passes through a hole in the anode *A*, is collimated as a fine pencil by the slit *S*, and in absence of any applied electric or magnetic fields strikes the phosphorescent tube wall at point *P*. However, when a magnetic field is applied across the electron path, the beam is deflected upward or downward from *P* according to the direction and strength of the field. If now an electric field of correct direction is superimposed over the magnetic field, it is possible to adjust the electric field strength until the displaced beam is returned to its original position at *P*. From the relations of the forces involved, the ratio of the electron charge *e* to its mass *m* may then be evaluated, and Thomson found thus *e/m* to

be 1.79×10^7 electromagnetic units per gram. Subsequent and more refined measurements have modified this value to

$$e/m = 1.758 \times 10^7 \text{ electromagnetic units/gram}$$
$$= 1.758 \times 10^8 \text{ coulombs/gram} \tag{1}$$

Charge on Electron. In order to resolve this ratio, the charge on the electron must be ascertained independently. First attempts at a measurement of e were made between 1897 and 1910 by Townsend, Thomson, Wilson, and Millikan. Although these studies did yield a concordant order of magnitude for e, they did not establish its value within the desired accuracy. It remained for Robert Andrew Millikan, through his famous oil-drop experiments in 1913, to determine the value of e with a high degree of precision.
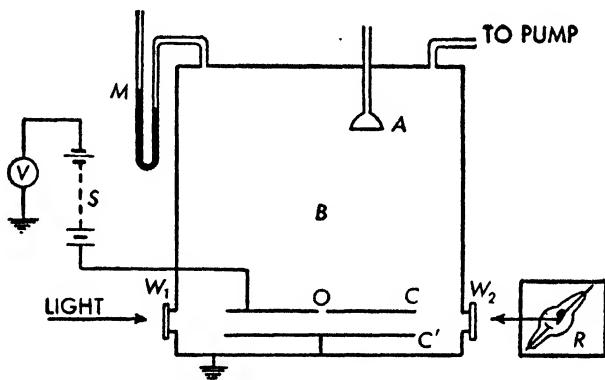
Fig. 2. Millikan's Oil-Drop Apparatus

A schematic diagram of Millikan's apparatus is shown in Fig. 2. It consisted of a chamber B, immersed in a thermostat (not shown) and filled with air whose pressure could be controlled by a vacuum pump and read on a manometer M. At the lower end of the chamber were installed two condenser plates C and C', across which could be applied a potential from the battery S. The space between the plates was illuminated by a light source located in front of a window in the chamber, W_1, and could be observed through a low-power telescope mounted in front of the apparatus and fitted with cross hairs a definite distance apart. By means of an atomizer A, a spray of fine droplets of oil was introduced into B. In due course one of these droplets worked its way down toward C and passed through an opening O in this plate into the space between C and C'. As soon as this was observed to happen, the opening O was closed, and the rate of fall of the droplet through the air under gravity was determined by measuring the time necessary for it to pass between

the cross hairs of the telescope. At this stage a beam of x rays from the source R was passed into the space between the plates through the window W_2. This beam caused the air molecules to become ionized, forming gaseous ions which would from time to time be captured by the oil droplet and impart to it a charge. When the electrostatic field was now turned on and applied from S across the condenser plates, an oil droplet with a charge upon it could be made to fall faster or actually rise against gravity with a velocity dependent on the direction of the field, its magnitude, and the charge on the drop. The rate of rise could readily be observed by timing again the rate of passage upward between the cross hairs. In fact, once a particle was trapped between the plates in this manner, it could be made to fall or rise at will by either shutting off or turning on the electrostatic field, and hence the measurements on the rates of rise and fall could be repeated on any given particle a number of times.

If we let v_1 be the rate of fall of the particle under gravity g, v_2 the rate of rise of the particle against gravity when an electrostatic field strength X is applied, and m' and e' the mass of and charge on the oil droplet, then in the oil drop experiments these quantities are connected by the relation

$$\frac{v_1}{v_2} = \frac{m'g}{Xe' - m'g} \qquad (2)$$

The values of m' required in equation (2) in order to find e' were obtained by use of a modified form of Stokes's law. From his many experiments with all types of drops and various field strengths Millikan found that e' was not constant. However, he did find that there was a *common factor* for all the values of e' which made the various charges observed *whole number multiples of the common factor*. This is an excellent confirmation of the atomicity of electricity, and indicates that the various oil droplets had captured and retained one, two, or more gaseous ions on their surfaces. The best value of this least common factor deduced by Millikan is $(4.774 \pm 0.005) \times 10^{-10}$ electrostatic units, and this must be, therefore, the charge of the unit of electricity, or the electron.

Since this work a redetermination of the viscosity of air and other corrections have shown that a better value of the electronic charge, e, is

$$e = 4.800 \times 10^{-10} \text{ electrostatic units/electron}$$
$$= 1.60 \times 10^{-19} \text{ coulombs/electron} \qquad (3)$$

The validity of this value of the electronic charge can readily be tested by using it to calculate Avogadro's number from the value of the faraday. Since a faraday represents the quantity of electricity associated with

Avogadro's number of unit electric charges, and since the quantity of electricity in unit electric charge is presumably e, N should be given by

$$N = \frac{\mathcal{F}}{e} = \frac{96,500 \text{ coulombs}}{1.60 \times 10^{-19} \text{ coulombs}}$$
$$= 6.03 \times 10^{23}$$

As a matter of fact, this procedure is one of the most precise means of evaluating N.

Mass of the Electron. From $e/m = 1.76 \times 10^8$ and $e = 1.60 \times 10^{-19}$ coulombs, the mass of the electron readily follows as

$$m = \frac{e}{(e/m)}$$
$$= \frac{1.60 \times 10^{-19}}{1.76 \times 10^8}$$
$$= 9.11 \times 10^{-28} \text{ g/electron} \qquad (4)$$

This mass of the electron may be compared with the mass of a hydrogen atom, which is, of course, the atomic weight of hydrogen divided by Avogadro's number, namely,

$$m_{\mathrm{H}} = \frac{1.008}{6.03 \times 10^{23}} = 1.67 \times 10^{-24} \text{ g}$$

From m_{H} and m it is seen that the mass of the hydrogen atom is 1833 times greater than the mass of the electron.

The mass of the electron given in equation (4) is the *rest mass*, i.e., the mass when the electron is either at rest or moving with velocities that are low compared to that of light. However, when the electron is moving at very high speeds, according to the theory of relativity the mass of the electron is increased in line with the equation

$$m = \frac{m_0}{\sqrt{1 - \left(\dfrac{v}{c}\right)^2}} \qquad (5)$$

where m_0 is the mass of the electron at rest, m the mass when the electron is moving with a velocity v, and c the velocity of light. From equation (5) it may be seen that m increases with v, until at $v = c$ the mass of the electron becomes infinite. Equation (5) is applicable not only to the electron, but also to any other body of mass m_0 at rest moving with extremely high velocities.

Wave Nature of the Electron. So far the electron has been treated as if it were only a particle. However, as was pointed out toward the end of Chapter IV, De Broglie in 1924 predicted that electrons should possess also wave properties, a prediction which was confirmed experi-

mentally by Davisson and Germer in 1925, and others. From this it must be concluded that electrons have the faculty of behaving both as particles and waves, depending on the manner in which they are observed. Thus, in a discharge tube the electrons exhibit their attributes as particles, whereas in electron diffraction experiments they act as electromagnetic waves analogous to x rays, and yield true diffraction patterns on reflection from crystalline surfaces. This duality of electronic behavior, exhibited also by light and even atoms, is an illustration of the *Heisenberg uncertainty principle* enunciated in 1927, which states that *it is impossible to define simultaneously the exact momentum and position of a body*. Precisely Heisenberg showed that the product of the uncertainty in the position of a body Δx, and the uncertainty in the momentum Δp, and hence the velocity, is even in the perfect experiment always equal to Planck's constant h; i.e.,

$$(\Delta x)(\Delta p) = h \qquad (6)$$

Consequently, as soon as any attempt is made to define exactly the position of a body, i.e., make Δx very small, Δp becomes large and leads to great uncertainty in Δp. Similarly any attempt to define Δp exactly leads to large uncertainties in Δx. Now, in studying the nature of the electron in electric discharge experiments we concentrate on the exact definition of the velocity and momentum of the electron, and arrive thus at the conclusion that the electron is a particle. On the other hand, when electron diffraction experiments are performed, the emphasis is on the position of the electron. Hence its momentum becomes an ambiguous quantity, and the electron behaves like a wave. From these considerations it may be deduced that the corpuscular nature of the electron becomes manifest when the particular experiment performed with the electron involves definition of its velocity, momentum, or energy, whereas the wave properties come to the fore when its position is being fixed. This definition of the conditions under which the electron acts as particle or wave is not an explanation; however, this is as far as we can go at present in elucidating this paradox.

The Proton. The researches described establish definitely that all matter contains negative electricity whose unit is the electron particle of relatively low mass. Since the atoms composing matter are generally electrically neutral, and since the mass of even the lightest atom, hydrogen, is very much greater than that of the electron, there must evidently be present in atoms positive electricity with which, in all probability, most of the mass is associated. The problem is, therefore, to find the unit of this positive charge and to ascertain its mass.

As a preliminary indication of the answer to this problem, let us consider in some detail the simplest and lightest of our atoms, hydrogen.

Under ordinary conditions this atom is electrically neutral and has a mass of 1.67×10^{-24} grams, which is 1833 times as great as that of the electron. When this atom is converted by ionization into an ion, it is found that its charge is exactly equal to that of the electron but is opposite in sign. Further, the mass of the positively charged hydrogen ion, which is called a *proton*, is essentially the same as that of the hydrogen atom. These facts would tend to suggest that the unit of positive electricity, complementary to that of the electron as negative electricity, is the hydrogen ion with a mass 1833 times that of the electron, and that the hydrogen atom is composed of one unit of positive and one of negative electricity, i.e., a proton and an electron.

It is not necessary to rely completely on this type of conjectural reasoning, for there are direct ways of producing positive rays and measuring their e/m ratio. In 1886 Goldstein discovered that there are present in a discharge tube not only negative cathode rays composed of electrons moving toward the anode, but also positive rays moving in the opposite direction. By perforating the cathode these *positive* or *canal rays* can be made to pass through the holes or canals of the cathode to the rear, and there studied by the effect of electric and magnetic fields upon them. By these methods it was established that these rays consist of *positively charged particles* which, unlike the electrons, have associated with them the mass of the atoms from which they come. Further, it was found that these particles are not emitted by the anode, but originate in the gas phase between the electrodes from the ionization of the gas atoms through electron bombardment. By appropriate means these rays can be generated also by emission of positive ions from anodes, in which instance the rays consist of ions of the metal used as anode. The significant fact revealed by all such studies is that, although unit positive charge may be associated with ions of various masses, no particle found is of a mass lower than that of the proton. Of added importance is the observation that the masses of all ions heavier than that of hydrogen are within close limits whole number multiples of the mass of the proton. These facts indicate that the positively charged hydrogen ion, or proton, is the unit of positive electricity, and that like the electrons the protons enter into the constitution of all atoms. This last conclusion is essentially a revival of Prout's hypothesis in more modern terms.

Positive Ray Analysis and Isotopes. J. J. Thomson's method of investigating positive rays was essentially similar to the one used in studying electrons. In this instance, however, the procedure proved to be rather difficult and not very precise. Although Dempster in 1918 developed a more satisfactory apparatus, it remained for Aston [1] to devise

---

[1] See Aston, "Mass Spectra and Isotopes," E. Arnold and Co., London, 1933.

and refine his *mass spectrograph* to such a point that masses of positively charged particles now can be determined with an accuracy of one part in 10,000 or better.

The principle of the mass spectrograph can be understood with the aid of the schematic diagram in Fig. 3. A beam of positive rays, narrowed by means of the slits S_1 and S_2, is passed between the two charged plates P_1 and P_2. Under the influence of the electrostatic field between
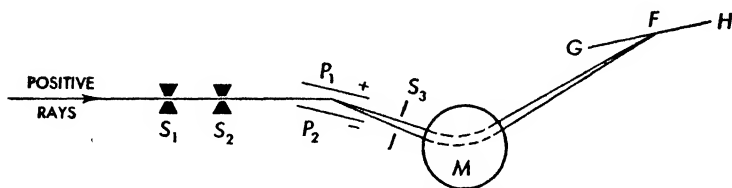
Fig. 3. Schematic Diagram of Aston's Mass Spectrograph

the plates the beam spreads out into a fan by downward deflection of the particles, the extent of deflection being determined by the charge and velocity of each particle. From the spread-out beam only those ions are selected which can pass through the narrow slit S_3, and these are sent into an electromagnetic field M directed to deflect the ions in a direction opposite to that of the electric field. From the relations obtaining for the apparatus it can be shown that for any given electric and magnetic field strength all charged particles of a given charge to mass ratio, even though not of the same velocity, will be brought to a focus at a point such as F. Furthermore, the focus of all particles of various charge to mass ratios will lie on a straight line situated along GH. Consequently, a photographic plate placed along GH will record by a narrow dark exposed band the point of focus of all the ions of a fixed charge to mass ratio, while a series of these dark bands will give the focused loci of all ions of various charge to mass ratios in the beam. The striking similarity of such an exposed plate to an ordinary spectrum plate accounts for the name *mass spectrograph* given to this apparatus.

Through careful study of the mass spectrograms of various elements, Aston was able to deduce the charge to mass ratios and therefrom the masses involved. In this manner he was able to arrive at very precise values of the atomic weights of the elements. The mass spectrograph also revealed that many elements, which from ordinary atomic weight determinations were considered to consist of atoms of the same mass, are in reality mixtures of atoms of different masses although of the same chemical properties. Such atoms of different mass and physical properties but with identical *chemical* properties are called *isotopes*. Thus chlorine, with a chemical atomic weight of 35.457, was shown to consist of two isotopes of atomic weights 35 and 37 on the O = 16.000 scale

mixed in such proportion as to yield an overall atomic weight of 35.457. Again, magnesium, chemical atomic weight 24.32, was shown to consist of isotopes of relative masses 24, 25, and 26. Whereas such elements as F, Na, Al, P, V, and I seem to possess no stable isotopes, elements such as He, Li, B, C, N, Cl, Co, Cu, Ga, Br, Rb, and Ag possess two isotopes, H, Be, O, Ne, Mg, Si, S, A, and K three, Ca, Cr, Fe, Sr, Ce, and Pb four, Ti, Ni, Zn, Ge, Zr, W, and Pt five, Se, Kr, and Pd six, Mo, Ru, Ba, Os, and Hg seven, Te eight, Cd and Xe nine, and Sn ten.   In every instance the abundance of each isotope in the mixture is such as to yield the overall chemically observed atomic weight.

In Table 1 are given the best available values for the atomic masses of a number of the lighter elements as deduced from the mass spectrograph. An inspection of this table reveals that to a very close approximation the masses of the various elements are very nearly whole numbers and multiples of the atomic weight of the hydrogen atom of mass 1.00813. This fact lends credence to the Prout hypothesis that all atoms are composed of protons as well as electrons. Again, the existence of isotopes accounts for fractional atomic weights deduced by ordinary chemical methods and shows why no satisfactory regularity could be observed in chemical atomic weights to substantiate Prout's suggestion.

TABLE 1

ATOMIC MASSES DEDUCED FROM MASS SPECTROGRAPH [1]

Atom	Atomic Number	Mass	Rounded Mass
H	1	1.00813	1
H	1	2.01473	2
He	2	4.00389	4
Li	3	7.0182	7
Be	4	9.0152	9
Be	4	10.0163	10
Be	4	11.0129	11
C	6	12.00398	12
C	6	13.0076	13
N	7	14.00750	14
N	7	15.0049	15
O	8	16.00000	16
O	8	18.0037	18
Si	14	27.9868	28
Si	14	28.9866	29
S	16	31.9823	32
Cl	17	34.9813	35
Cl	17	36.9788	37

For strict applicability of Prout's hypothesis it would be required that the mass of various atoms should be exact whole number multiples of the mass of the hydrogen atom, 1.00813. Actually, however, this is

[1] Livingston and Bethe, Rev. Mod. Physics, **9**, 366 (1937).

not the case. Thus helium, with four protons, should have a nuclear mass of $4 \times 1.00813 = 4.03252$, but the actually observed mass is 4.00389. This mass loss of 0.02863 gram in the formation of a helium atom from four hydrogen atoms can be accounted for in terms of an equation derived by Einstein, in which he shows that mass may be converted to energy, or vice versa, according to the relation

$$E = mc^2 \tag{7}$$

In this equation E is the energy in ergs, m the mass in grams, and c the velocity of light, 3×10^{10} cm/second. If it is postulated now that the formation of the helium atom from four hydrogen atoms takes place with an evolution of energy equivalent to the loss in mass, then the process in question involves a liberation of

$$\begin{aligned} E &= 0.0286 \times (3 \times 10^{10})^2 \\ &= 2.57 \times 10^{19} \text{ ergs} \\ &= 6.1 \times 10^{11} \text{ cal} \end{aligned}$$

In a like manner may be explained the mass defects observed in other atomic nuclei.

Separation of Isotopes. Not only may the existence of isotopes be demonstrated, but the isotopes may be separated by a number of methods. Since the isotopes of an element are indistinguishable from each other in chemical behavior, all methods for isotope separation must depend on slight physical differences among them.

Among the methods which have been employed to separate or enrich isotopes are the mass spectrograph, diffusion, centrifuging, thermal diffusion, electrolysis, fractional distillation, and chemical exchange. It is not possible to describe here either these methods or the results obtained. This may be found elsewhere.[1] Nevertheless, it is of considerable interest to point out the differences found between ordinary hydrogen and its heavier isotope *deuterium*, symbol D. Deuterium was discovered spectroscopically in 1932 by Urey and his coworkers at Columbia University. In ordinary hydrogen gas deuterium is present as about one part in 6400 parts of light hydrogen. The heavier isotope is usually concentrated by electrolysis of aqueous alkali solutions, in which instance the lighter hydrogen atoms escape more readily than the heavier deuterium, and the solution becomes more concentrated in D_2O, *heavy water*. By repeating the electrolysis a number of times, pure D_2O can be prepared and studied.

In Table 2 are contrasted some physical properties of hydrogen and deuterium, while in Table 3 are compared the properties of water and

[1] See Aston, *loc. cit.*; also Taylor and Glasstone, "Treatise on Physical Chemistry." D. Van Nostrand Company, Inc., New York, 1942, Vol. I, pp. 54–76.

TABLE 2

COMPARISON OF PROPERTIES OF HYDROGEN ISOTOPES

Property	H_2	D_2
Molar volume of solid	26.15 cc	23.17 cc
Triple point	13.92° K	18.58° K
Heat of fusion	28 cal/mole	47 cal/mole
Boiling point	20.38° K	23.50° K
Heat of vaporization at triple point	217.7 cal/mole	303.1 cal/mole

heavy water. It will be observed that the differences in the properties of pure hydrogen and deuterium are fairly pronounced, and that these differences are carried over into the compounds formed from these. The differences in the two species of water molecules, H_2O and D_2O, affect also the solubility, conductance, equilibrium relations, and the rates of various reactions which may be conducted in these media as solvents. In this connection it may be mentioned that the physical differences between hydrogen and deuterium are much more pronounced than are those between isotopes of other elements where the relative difference in mass is not anywhere near as great as it is here.

TABLE 3

COMPARISON OF PROPERTIES OF H_2O AND D_2O

Property	H_2O	D_2O
Density, 25°/25°	1.000	1.1079
Melting point	0.00° C	3.82° C
Boiling point	100.00° C	101.42° C
Surface tension, 20° C	72.75 dynes/centimeter	67.8 dynes/centimeter
Temperature of maximum density	4.0° C	11.6° C
Dielectric constant, 25° C	78.54	78.25
Viscosity, 20° C	10.09 millipoise	12.6 millipoise
Heat of fusion	1436 cal/mole	1510 cal/mole
Heat of vaporization	10,480 cal/mole	10,740 cal/mole

Earlier Theories of Atomic Structure. Although the preceding evidence points to protons and electrons as the constituents of all matter, it does not answer the question of "atomic architecture," namely, how the protons and electrons are arranged to yield the various atoms. An early attempt to answer this question was made in 1903 by Lenard, but his simple theory was displaced in 1911 by Rutherford's proposal of a nuclear atom. Rutherford's theory is based on some observations made by Geiger and Marsden and by himself on the scattering of α particles by metallic foils. α particles are emission products of certain radioactive disintegrations. They have a mass approximately four times that of the proton and bear two positive charges. These particles are in fact helium nuclei, i.e., helium atoms from which two electrons have been removed. When Geiger and Marsden bombarded thin metallic foils

with these α particles, they found that whereas most of these went through the foils unaffected, about one particle in 20,000 suffered a violent deflection through angles of 90° or greater. In seeking an explanation of these results it must be realized that electrons, because of their low mass, would not be able to deflect α particles. Again, to cause the sharp deflection of the particles through the large observed angles would require strong forces which can arise only from interaction of particles of the same charge as those used in the bombardment. Finally, since most of the α particles go through unaffected, these sources of deflection cannot be continuous throughout the foil, but must be localized in
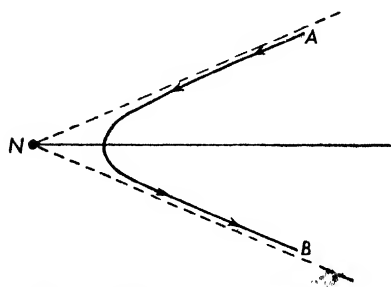
spaces that are relatively small compared to the space occupied by the foil as a whole. Basing his argument on these points, Rutherford suggested that atoms consist of miniature solar systems in which all of the positive charge is located in a nucleus at the center of the atom, while the electrons required for electrical balance revolve about the nucleus at some distance from it, much as the planets revolve

Fig. 4. Scattering of α Particles by Nuclei

about the sun. As in terms of this picture the atom consists mostly of empty space, α particles directed against a thin metallic foil should be able to pass through it without deflection as long as they do not approach too near to a positive nucleus. However, when an α particle is so directed as to come near the nucleus, the strong repulsive force developed between the nucleus and the positive α particle causes the latter to be deflected in the manner indicated in Fig. 4. Here N is the nucleus of an atom, while AB is the path of the α particle on approach and deflection.

From a careful mathematical analysis of the scattering Rutherford was able to deduce that the charge on the nucleus responsible for the observed deflections was equal approximately to one-half the atomic weight of the metal constituting the foil. Further, he was able to estimate that the dimensions of the nuclei were of the order of $10^{-12} - 10^{-13}$ cm, dimensions which when considered along with the volume of the electrons involved are only $10^{-12} - 10^{-15}$ of the volumes actually occupied by the atoms. From these figures it follows that only $1/10^{12}$ to $1/10^{15}$ of the volume of an atom is occupied by protons and electrons, the rest being nothing but empty space.

The Charge on the Nucleus. Rutherford's earlier researches on α-particle scattering, although of fundamental importance, did not lead

him to an exact definition of the nuclear charge. The first precise statement that the *nuclear charge of an atom is identical with its atomic number* was made by van der Broeck in 1913. None the less, the first direct evidence on this point is due to Moseley's investigations (1913, 1914) on the x-radiations emitted by various elements. These will now be described.

The student will recall from Chapter IV that x rays are emitted when a metallic target is bombarded with rapidly moving electrons. No x-radiation from a given target is observed until the electrons acquire a certain threshold speed which is different and characteristic for each metal. Once the required speed of bombarding electrons is attained or exceeded, it is observed that each metal emits x-radiation which is not continuous, but consists of series of lines possessing fixed frequencies characteristic of the target substance. The various sets of lines thus obtained can be divided into series designated in turn as the *K*, *L*, *M*, etc., series, while the lines within each series are referred to in sequence as the α, β, γ, etc., lines. Thus, the K_α line is the first line in the K series, the L_β is the second line in the second series, etc.

Moseley undertook the task of studying the x-ray spectra of elements lying between aluminum and gold in the periodic table to ascertain whether any regularities exist among them. As a result of this exhaustive investigation he found that the spectra observed followed the same sequence as do the elements in the periodic table. And, what is more important, he also found that, if we take the frequencies of any particular line in all the elements, such as the K_α or K_β lines, then the frequencies are all related to each other by the equation

$$\sqrt{\nu} = a(Z - k) \tag{8}$$

where ν is the frequency of any particular K_α or K_β, etc., line, Z is the *atomic number* of the element in the periodic table, and a and k are constants for any particular type of line. The regularity thus observed can be obtained only when the *atomic number* is used, but not with the atomic weight. This suggests that the fundamental quantity involved in x-radiation is not the atomic weight but the atomic number. Furthermore, since in all probability x-radiation arises in energy shifts taking place in extranuclear electrons, and since x-radiation of the various elements exhibits regularity with atomic number, Moseley concluded that the change in atomic number from element to element represents the regular variation in the number of extranuclear electrons in the neutral atom; i.e., *the number of extranuclear electrons is identical with the atomic number of an element.* Moreover, since the number of extranuclear electrons in an atom must be equal to the charge of the nucleus for overall electroneutrality, it must follow also that *the atomic number gives the*

number of units of positive charge on the nucleus. The correctness of this conclusion has subsequently been verified by Chadwick's work on scattering of α particles, and by certain observations on radioactive disintegrations.

Distribution of Electrons and Protons in an Atom. We are prepared now to ascertain the number of electrons and protons in an atom and their distribution. On the basis of the oxygen scale of atomic weights, the mass of the proton is essentially unity. Further, because of the relatively light mass of electrons compared to that of protons, the mass of atoms will be due primarily to the latter. Therefore, if we let A be the atomic weight of any given specie of atom, this will be also the number of protons; and, again, all of these will be localized in the nucleus. To give the atom as a whole no net electric charge, there must evidently be present in each atom also A electrons. Now, since Moseley showed that only Z of these, where Z is the atomic number, are present outside the nucleus, the number of electrons present in the nucleus along with the protons must be $(A - Z)$.

The following, then, is the picture of an atom. Each atom consists of a nucleus composed of A protons and $(A - Z)$ electrons, giving thereby the nucleus a net charge Z. To balance this nuclear charge Z electrons are distributed outside the nucleus and revolve about it as a center. Thus, in the hydrogen atom of unit mass and $Z = 1$ there must be only one proton in the nucleus and one electron revolving about it. Again, in beryllium with $A = 9$ and $Z = 4$ there must be a nucleus composed of nine protons and five electrons, and four external electrons. Finally, in an atom such as uranium, with $A = 238$ and $Z = 92$, there must be present 238 protons and 146 electrons in the nucleus, and 92 external satellite electrons.

Rutherford believed that the external electrons may occupy any and all positions outside the nucleus and possess thereby energies which can vary continuously. However, this concept of continuous variation in energy is contradicted by atomic spectra, which are not continuous but discontinuous. Again, since a revolving electron is a charged body in motion, then according to classical electrodynamics such a body should radiate energy continuously. As a result of this loss of energy by radiation the orbit of revolution of an electron should get smaller and smaller, until eventually the electron should fall into the nucleus and be retained there. Such behavior on the part of electrons has never been observed. To overcome these difficulties inherent in the Rutherford atom Niels Bohr advanced in 1913 his now famous theory of atomic structure. Before this theory can be presented, however, two other subjects must be introduced first, namely, Planck's quantum theory of radiation and the emission of line spectra by excited atoms.

The Quantum Theory of Radiation. When radiation strikes any surface of a body, part of the radiant energy is generally reflected, part is absorbed, and part is transmitted. The reason for the incomplete absorption is that ordinary bodies are as a rule imperfect absorbers of radiation. In contrast to ordinary bodies we have the *black body*, which by definition is the perfect absorber of energy, and retains any radiant energy that strikes it without reflection or transmission. Although a blackened metallic surface or carbon black approximates fairly closely to a black body, a hollow sphere, blackened on the inside and with a small opening, meets the defined condition much more satisfactorily. Any radiation that enters through the small opening is reflected repeatedly from the walls of the enclosure until all of the energy eventually becomes absorbed.

It is a characteristic property of any black body that it is not only a perfect absorber of radiant energy, but also a perfect *radiator*. In fact, of all bodies heated to a given temperature, it is the black body which will radiate the maximum amount of energy possible for the given temperature. Furthermore, a black body is in thermal equilibrium with its surroundings, and radiates in any given time per unit area the same amount of energy as it absorbs.

The total amount of energy E radiated by a black body per unit area and time is given by the *Stefan-Boltzmann fourth power law*, namely,

$$E = \sigma T^4 \tag{9}$$

where T is the absolute temperature, while σ is a universal constant equal to 5.673×10^{-5} for energy in ergs, time in seconds, and area in square centimeters. This energy is not emitted with a single frequency of radiation, nor is it uniformly distributed along the spectrum. Lummer and Pringsheim showed that the energy emitted depends on the temperature and wave length in the manner indicated in Fig. 5. For each temperature there is a wave length at which the energy radiated is a maximum. Again, the position of this maximum shifts continually with increase in temperature toward the lower wave lengths, and the maximum itself is the more pronounced the higher the temperature.

Wien in 1896 and again Lord Rayleigh in 1900 made an attempt to account for this spectral distribution of black body radiation on the basis of classical concepts of *continuous emission* of radiation. Although Wien's equation proved fairly satisfactory at low wave lengths and Rayleigh's equation at high wave lengths, neither theory was sufficient to account completely for the observed phenomena. This failure of classical theories of radiation led Max Planck in 1900 to discard these theories, and to come forward with the bold hypothesis that *black bodies*

radiate energy not continuously, but discontinuously in energy packets, called quanta, given by the relation

$$E = h\nu \tag{10}$$

where E is the quantum of energy radiated, ν is the frequency, and h, called *Planck's constant*, is a universal constant equal to 6.624×10^{-27} erg-sec. Equation (10) is the fundamental relation of the *quantum theory of radiation*. Planck considered the black body to consist of
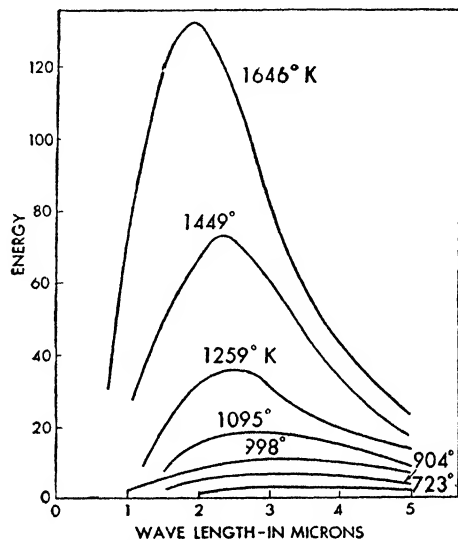
Fig. 5. Distribution of Energy in Black Body Radiation

oscillators of molecular dimensions, each with a fundamental vibration frequency ν, and that each oscillator could emit energy either in the unit quantum $E_1 = h\nu$, or in *whole number multiples* n thereof, $E_n = nh\nu$. On this basis he was able to deduce, for the energy E_ν radiated by a black body at any frequency ν, the relation

$$E_\nu = \frac{2\pi\nu^2}{c^2} \cdot \frac{h\nu}{e^{h\nu/kT} - 1} \tag{11}$$

where c is the velocity of light, T the absolute temperature, and k the gas constant per molecule, i.e., R/N. This equation not only reproduced in excellent fashion the spectral distribution of energy, but reduced as well for low wave lengths to Wien's equation and for high wave lengths to the Rayleigh equation.

Planck's success with the quantization of black body radiation led Albert Einstein in 1905 to a generalization of the quantum theory. Einstein postulated that not only is black body radiation quantized,

but so is all other absorption or emission of radiant energy; i.e., *any radiation energy absorbed by a body must be absorbed in whole units or quanta whose magnitude depends on the frequency according to equation (10), and* similarly *all energy is radiated by bodies in quanta or multiples thereof.* Furthermore, Einstein argued that radiation is not only emitted or absorbed in quanta, but is also propagated through space in these units, called *photons* of light. This theory definitely ascribes to light a corpuscular nature, just as the researches described in the earlier parts of the chapter established the corpuscular nature of electricity and matter.

Einstein's theory of the corpuscular nature of all light has been amply substantiated by observations on photoelectric electron emission by light, by the inverse photoelectric effect, i.e., the emission of radiation by electron bombardment, as in the generation of x rays, by the Compton effect arising from collisions between photons and electrons, and several other phenomena. Although details of these cannot be given here, they may readily be found in books on the quantum theory.[1]

The Line Spectra of Atoms. We have seen that bodies on being heated to high temperatures emit radiation. This radiation can be passed through a spectrograph in which the radiation can be resolved into its component wave lengths and recorded photographically. When such spectrograms are analyzed, it is found that solids generally yield spectra that are continuous. On the other hand, gases and vapors yield instead a series of lines, called *line spectra*, or bands, called *band spectra*, which consist of many lines close together. The *band spectra* are radiations emitted by *molecules*, whereas the *line spectra* are due to *atoms*. At present we are interested only in the latter, and so we shall turn to a discussion of these.

Of all the elements the simplest line spectrum is that exhibited by atomic hydrogen. The spectrum of this element consists of a number of lines which can be classified into groups or series. Each group or series of lines is related by a series formula, which for the *Balmer* series, appearing in the visible spectral range, takes the form

$$\frac{1}{\lambda} = \bar{\nu} = R_\mathrm{H}\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \tag{12}$$

Here λ is the wave length of the line, whose reciprocal, $\bar{\nu}$, is called the *wave number*, i.e., the number of waves per centimeter, R_H is a constant, called the *Rydberg constant* and equal to 109,677.76 cm^{-1}, and n is a running number taking on for the various lines of this series values of 3, 4, 5, etc. Similarly the lines in the various other series found in hydro-

---

[1] See especially S. Dushman's chapter in Taylor and Glasstone, *op. cit.*, Vol. I, pp. 192-200.

gen can be expressed by formulas analogous to equation (12). In fact, the lines of all series can be represented by the general expression

$$\bar{\nu} = R_\text{H} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \tag{13}$$

where the values of n_1 and n_2 for the various series are summarized in Table 4. Column 4 of this table lists also the spectral regions in which the lines of the particular series are found. The significant fact to be observed here is that every line in any given series can be represented as a difference of two terms, one R_H/n_1^2, where n_1 has a fixed value for a given series, and a second R_H/n_2^2, where n_2 can take on a series of integral consecutive values beginning with $n_2 = n_1 + 1$. The significance of this important regularity will appear as soon as the Bohr theory is presented.

TABLE 4

SPECTRAL SERIES OBSERVED IN ATOMIC HYDROGEN

Series	n_1	n_2	Spectral Region
Lyman	1	2, 3, 4, $\cdots$	Ultraviolet
Balmer	2	3, 4, 5, $\cdots$	Visible
Paschen	3	4, 5, 6, $\cdots$	Infrared
Brackett	4	5, 6, 7, $\cdots$	Infrared
Pfund	5	6, 7, $\cdots$	Infrared

The spectra of elements other than hydrogen are as a rule much more complicated. The spectra of ionized gaseous atoms such as singly ionized helium, He^+, doubly ionized lithium, Li^{++}, or trebly ionized beryllium, Be^{+++}, bear a striking resemblance to that of hydrogen, and these are said to be hydrogenic in character. With the heavier atoms, however, and with the unionized atoms of the elements mentioned the situation is more complex, and a different method of classification is necessary. Nevertheless, the spectra of all these elements can be arranged into series; and every series can again be represented as a difference of two terms, one fixed and characteristic of the series, another integrally variable with each line.

Bohr's Theory of the Hydrogen Atom. The energies emitted by excited atoms with frequencies corresponding to those of the line spectra observed must evidently have their source in changes occurring within the atoms. To account for the line spectra of elements and to circumvent the objections leveled against the Rutherford atom, Niels Bohr advanced in 1913 a theory of atomic structure radically different from any that preceded it. Along with Rutherford, Bohr considered the atom to consist of a nucleus with electrons revolving about it. However, whereas Rutherford's theory permitted the electron to revolve in any

and all possible orbits, Bohr postulated that the only possible orbits for
electron revolution are those for which the angular momentum is a
whole number multiple n of the quantity $(h/2\,\pi)$, h being Planck's con-
stant. This assumption constitutes a *quantization of the angular momen-
tum* of the electron. Bohr further postulated that as long as any given
electron stays in a given orbit it does not radiate energy, despite the
demands of classical electrodynamics, and hence the energy of an elec-
tron remains constant as long as it does not change orbits. This second
postulate introduces the concept that there are definite *energy levels* or
stationary states within the atom in which an electron possesses a definite
and invariable energy content. Finally Bohr assumed as his third pos-
tulate that each line observed in the spectrum of an element results from
the passage of an electron from an orbit in which the energy is E_2 to
one of lower energy E_1, and that this difference in energy is emitted as
a quantum of radiation of frequency ν in line with the Planck equa-
tion (10), namely,

$$\Delta E = E_2 - E_1 = h\nu \qquad (14)$$

Bohr proceeded immediately to apply these ideas to the hydrogen
atom. If the single electron in hydrogen revolves about the proton
nucleus in a circular orbit, the situation in the hydrogen atom can be
represented essentially by the diagram shown in
Fig. 6. In this diagram point A represents the
proton of positive charge $+ e$, where e is the elec-
tronic charge. About this nucleus in an orbit of
radius r revolves an electron B, with a tangential
velocity v, charge e equal but opposite to that of
the proton, and a mass m. This electron is
attracted on the one hand by electrostatic forces
to the nucleus, while at the same time it experi-
ences a centrifugal force which tends to make it

Fig. 6. Bohr's Model of
Hydrogen Atom

escape from its orbit. For the electron to remain in its orbit these two
forces must be equal. Equating them, and introducing his postulate of
quantization of angular momentum, Bohr was able to show that the
velocity of the electron in its orbit should be given by

$$v = \frac{2\,\pi e^2}{nh} \qquad (15)$$

and its energy E by

$$E = \frac{-\,2\,\pi^2 m e^4}{n^2 h^2} \qquad (16)$$

Equation (16) gives the total energy of an electron in a hydrogen
atom. It will be observed that this equation involves only known con-

stants and the integer n which may have only integral values 1, 2, 3, etc. For any given value of n, which is called the *principal quantum number*, the energy of the electron will have a definite value, and this will be the constant energy of the particular energy level or stationary state. The energies of the various energy levels, since they are proportional to $1/n^2$, will decrease according to the square of the quantum number, so that the level for which $n = 2$ will have only 1/4 the energy of the level for which $n = 1$, the one with $n = 3$ 1/9 the energy, etc. In other words, the energy levels which an electron in a hydrogen atom can occupy do not change continuously, but must involve abrupt changes in energy on passage from one level to another.

This same abruptness in change is exhibited also by the radii of the paths in which the electron revolves, which are given by

$$r = \frac{n^2h^2}{4\pi mc^2} = (0.53 \times 10^{-8})n^2 \text{ cm} \qquad (17)$$

We see, therefore, that the orbit closest to the nucleus, i.e., for $n = 1$, has a radius of 0.53 Å, the one for $n = 2$ a radius of 4×0.53 Å, the one for $n = 3$, 9×0.53 Å, etc. The relative locations of these orbits for several values of n are illustrated in Fig. 7.

Bohr's Theory of the Hydrogen Spectra. According to Bohr's third postulate, spectral lines are supposed to result from electron jumps from one energy level to another, and as a result of the energy decrease a quantum of energy is emitted with a frequency given by equation (14). This point can be tested very readily. Suppose the electron is originally in an orbit for which $n = n_2$ and that it changes into an orbit for which $n = n_1$. In its initial energy state the electron will have, according to equation (16), the energy

Fig. 7. Bohr Orbits for Electrons in Hydrogen Atom

$$E_{n_2} = -\frac{2\pi^2me^4}{h^2n_2^2}$$

while in the final state

$$E_{n_1} = -\frac{2\pi^2me^4}{h^2n_1^2}$$

The difference in energy, $\Delta E = E_{n_2} - E_{n_1}$, should be emitted then as a quantum of energy $h\nu$. Hence the wave number $\bar{\nu}$ of the emitted line should be given by

$$h\nu = \Delta E = E_{n_2} - E_{n_1}$$
$$= -\frac{2\,\pi^2 me^4}{h^2 n_2^2} + \frac{2\,\pi^2 me^4}{h^2 n_1^2}$$
$$\nu = \frac{2\,\pi^2 me^4}{h^3}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$
$$\bar{\nu} = \frac{\nu}{c} = \frac{2\,\pi^2 me^4}{h^3 c}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \tag{18}$$

The striking similarity of equation (18) to the general equation (13) for the spectral lines of hydrogen is immediately evident; and, if Bohr's theory is correct, the Rydberg constant R_H should be equal to

$$R_H = \frac{2\,\pi^2 me^4}{h^3 c} \tag{19}$$

Substitution into equation (19) of the available values of m, e, h, and c yields for R_H 110,000 cm^{-1}, as against the experimental value R_H = 109,678 cm^{-1}. This remarkably close agreement constitutes a strong confirmation of Bohr's postulates and his explanation of the origin of the hydrogen spectrum.

Further comparison of equations (18) and (13) supplies the significance of the various series observed in the spectral lines emitted by the hydrogen atom. In the Lyman series $n_1 = 1$, and hence the lines observed result from passage of electrons from the second, third, fourth, etc., orbits into the first. Similarly the Balmer series, for which $n_1 = 2$, arise from passage of electrons from the third, fourth, etc., orbits into the second orbit. Likewise the Paschen, Brackett, and Pfund series arise from electron transitions from higher orbits into the third, fourth, and fifth orbits respectively. These various transitions for the sundry series are illustrated schematically in the energy level diagram for the hydrogen atom shown in Fig. 8. Each horizontal line in this figure illustrates the energy content of a particular level corresponding to the indicated quantum number, while the vertical lines represent the various electron transitions from which arise the spectral lines observed in each series.

The Bohr theory of the hydrogen spectrum can be extended also to account for the *hydrogen-like* spectra of singly ionized helium, He$^+$, doubly ionized lithium, Li^{++}, and trebly ionized beryllium, Be^{+++}. By a hydrogen-like spectrum is meant one that results from a system consisting of a nucleus and a single external electron. In the instances mentioned the wave numbers of the respective lines are given by

$$\bar{\nu} = \left(\frac{M}{m + M}\right) R Z^2 \left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \tag{20}$$

where M is the mass of the respective nucleus, Z is its charge, m is the mass of the electron, and R is a constant. Here again the agreement between theory and experiment is found to be quite satisfactory.
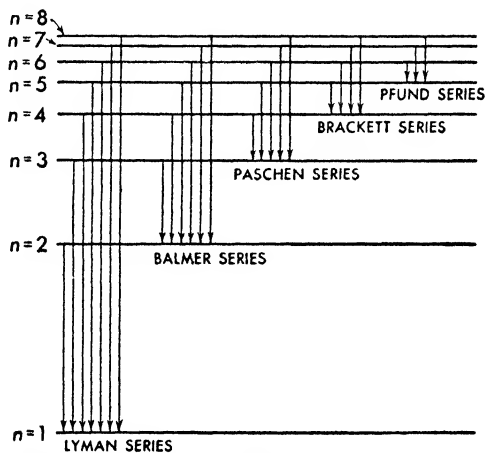
Fig. 8. Origin of Spectral Series in Hydrogen

Sommerfeld Extension of Bohr Theory. In describing the spectra of atoms it was assumed that each atom on excitation emits a series of lines each of which is a single unit. Actually, however, it is found on strong resolution that many lines are not single, but consist of two or more lines spaced so closely together as to constitute apparently a single line. This appearance of multiple fine lines in place of a single line is generally referred to as *multiplet* or *fine structure* of spectral lines.

Now, although the Bohr theory does account for the positions of the various lines in hydrogen or hydrogenic spectra, it does not explain their fine structure. To meet this deficiency Arnold Sommerfeld advanced in 1915 the suggestion that electrons need not revolve necessarily in circular orbits, but that these may be elliptical. Assuming, in line with all elliptical motions, that the nucleus is situated at one of the foci of the ellipse, and resolving the angular momentum along a radius vector and a direction perpendicular to it, Sommerfeld showed that, on the basis of quantization of both resolved momenta, the energy of the hydrogen atom should be given by

$$E = -\frac{2\,\pi^2 e^4 m}{h^2(n_r + k)^2} \tag{21}$$

In this equation n_r is an integer called the *radial quantum number*, while k is an integer called the *azimuthal quantum number*. Comparison of this

equation with the Bohr equation (16) shows that the two are identical on setting the principal quantum number n equal to

$$n = n_r + k \qquad (22)$$

Further, it can be shown that the principal and azimuthal quantum numbers are related to the lengths of the major and minor axes of the ellipse by the equation

$$\frac{n}{k} = \frac{\text{Length of major axis}}{\text{Length of minor axis}} \qquad (23)$$

From equation (23) it may be seen that when $n = k$, i.e., when $n_r = 0$, the two axes are the same and the path is therefore a circle. When k is less than n, however, various elliptical paths are possible depending on the values of k. These values of k may be $(n - 1)$, $(n - 2)$, etc.,

down to $k = 1$, since when $k = 0$ the ellipse would degenerate into a straight line through the nucleus — an apparently impossible situation. A series of the possible paths for $n = 4$ and the values of k resulting therefrom, namely, $k = 4$, $k = 3$, $k = 2$, and $k = 1$, is given in Fig. 9.

With his theory Sommerfeld was able to account not only for the fine structure of hydrogen spectra, but also for those of
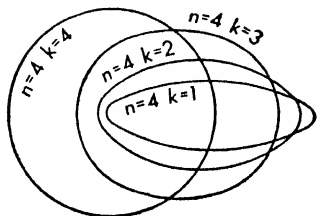
Fig. 9. Electron Orbits in Hydrogen (Sommerfeld)

the hydrogenic type. Nevertheless, there is an objection which can be leveled against this theory. Although it does give the correct value of n and the correct number of values that k may adopt, the actual values of the k's are wrong. Thus spectroscopic studies show that k values should not vary from $k = n$ down to $k = 1$, but from $k = n - 1$ down to $k = 0$. Yet the Bohr theory, as modified by Sommerfeld, is believed to be essentially correct in its pictorial aspects of the electron behavior in atoms, and is still the simplest working model of atomic architecture.

Before leaving the Bohr-Sommerfeld theory of atomic structure, it may be well to summarize its physical significance. According to this theory each atom is a miniature "solar system" in which the role of the sun is taken by the positive nucleus, while the role of the planets is taken by the external electrons. These electrons revolve about the nucleus in orbits which may be circular or elliptical. The orbits of revolution possible are not any and all imaginable, but only those permitted by the quantization of the angular momenta of the electrons. In these various orbits the electrons possess definite energies, and this fact leads us to the concept of stationary or energy levels. However, whereas Bohr considered each energy level to be constituted of only one definite energy,

the Sommerfeld extension points to the fact that any energy level corresponding to a given value of the principal quantum number n may in fact be composed of a series of *sublevels*, each with a different value of the azimuthal quantum number, and each only slightly different from the other in total energy. Finally, the spectral lines emitted by atoms result from the passage of electrons from one energy level to a lower one. Were there no sublevels present, each line thus emitted would be a singlet with no fine structure. But, since sublevels may be present, transitions from those in any one major level to others at lower levels lead to a number of closely spaced lines instead of only one line, and this, therefore, is the cause of the appearance of fine structure.

Quantum Numbers. The Bohr-Sommerfeld theory indicates that two quantum numbers, the principal quantum number n and the azimuthal quantum number k, are necessary to express the energy of an electron in an atom. Spectroscopic study of atoms in magnetic fields shows, however, that in actuality *four* quantum numbers are required: (1) the *principal quantum number n*, (2) the *azimuthal quantum number*, which in its correct form is represented by l, (3) the *magnetic quantum number m*, and (4) the *spin quantum number s*.

The principal quantum number n of any electron in any atom represents the *major energy level* to which an electron belongs. This number may take on the integer values 1, 2, 3, etc., depending on whether the electron is in the first, second, third, etc., major level. Corresponding to every major energy level n, there are n *sublevels* possible whose designation is given by the azimuthal quantum number l. This quantum number may have values ranging from $l = 0$ up to $l = (n - 1)$. For instance, for any major level of $n = 1$ there is only one value of l possible, $l = 0$. Again, for a level $n = 2$ we may have $l = 0$ and $l = 1$, whereas for a value of $n = 4$ we may have $l = 0, l = 1, l = 2,$ and $l = 3$. In any instance the total number of l's possible is equal to the principal quantum number n.

The magnetic quantum number arises from the fact that in a strong magnetic field *each* spectral line can be split into a number of component lines. This effect, called the *Zeeman effect* after its discoverer, can be explained on the supposition that each sublevel can exist in fact as a number of closely related levels whose presence is brought out in a magnetic field. For the description of these extra levels another quantum number m, called the *magnetic quantum number,* is required. The values which this quantum number can have are determined by the values of l of the sublevels, and range from $- l \longrightarrow 0 \longrightarrow + l$, i.e., a total of $(2 l + 1)$ values for each value of l. For example, for $l = 0$ there is only one value of m, namely, $m = 0$. Again, for $l = 1$ we may have $m = - 1$, $m = 0$, and $m = + 1$, a total of three values. On the other hand, if

$l = 3$, we have seven values of m, these being $m = -3, -2, -1, 0, +1, +2, +3$.

Finally, the *spin quantum number* s arises from a suggestion made by Uhlenbeck and Goudsmit in 1925 that an electron in its motion in an orbit may rotate or spin about its own axis. Such a spin would contribute to the angular momentum of the electron, and would modify thus the energy relations. Assuming that this spin is also quantized, there are only two possible values which s may have, $s = +1/2$ or $s = -1/2$, depending on whether the electron spins in one direction or in another.

The four quantum numbers and the values they may take on may be summarized as follows:

1. *Principal quantum number, n.* This number can only have the integral values $n = 1, 2, 3$, etc.

2. *Azimuthal quantum number, l.* For each value of n there may be n values of l, namely, $l = 0, 1, 2 \cdots (n - 1)$

3. *Magnetic quantum number, m.* For each value of l there may be $(2l + 1)$ values of m. These range from $m = -l$ through $m = 0$ to $m = +l$ in integral values; i.e., $m = 0, \pm 1, \pm 2 \cdots \pm l$.

4. *Spin quantum number, s.* There are only two possible values of s, $s = 1/2$ and $s = -1/2$ for each value of l.

By specifying its four quantum numbers, the "address" of any given electron in a given atom is completely defined; i.e., the four quantum numbers definitely locate each electron in the major energy level (n), the particular sublevel (l), the sub-sublevel (m), and the direction of its spin (s).

Wave Mechanical Theory of Atomic Structure. In conjunction with the discussion of the electron (p. 671) and of electron diffraction (p. 138) it was pointed out that according to De Broglie every moving particle can be considered to have associated with it a wave with a frequency determined by its velocity. Starting with this idea, Schrödinger in 1926 developed the theory of *wave mechanics*, and showed that on the basis of this theory any moving particle must obey the fundamental wave equation

$$\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0 \tag{24}$$

In this Schrödinger wave equation m, E, and V are, respectively, the mass, total energy, and potential energy of the moving body, h is again Planck's constant, ψ is a quantity called the *wave function* of the moving body, and x, y, and z are the three perpendicular coordinates along which the motion of the body may be resolved. The wave function is a quantity rather difficult to describe here. Suffice it to say that for an electron the value of ψ^2 at any point may be interpreted as the density

of the electric charge at the given point. Further, the wave function ψ has the mathematical attributes of being finite, single valued, and continuous for any physical situation.

Solution of equation (24) is not a simple task; however, it has been carried through for the hydrogen atom. This shows that the only way in which equation (24) can be valid for this atom is for the total energy of the electron to be given by

$$E = -\frac{2\,\pi^2 m e^4}{n^2 h^2} \tag{25}$$

where n is an integer. This equation is identical with equation (16) for the hydrogen atom deduced by Bohr. However, whereas derivation of equation (16) required on Bohr's part the arbitrary assumption of the quantization of the angular momentum, equation (25) follows directly from the mathematical requirements of the wave equation. Furthermore, the solution also indicates that for any given value of n there must be several values of the azimuthal quantum number such that $l = 0, 1, 2 \cdots (n-1)$, and that there is also a magnetic quantum number such that, for every value of l, $m = 0, \pm 1, \pm 2, \cdots \pm l$. Finally, the spin requirements can also be deduced from this theory. We have here, therefore, a theory that ties together all the various factors introduced on postulate and experiment, and yields these as essential and integral requirements of the mathematical relations involved.

It is true that in doing this service the theory of wave mechanics removes a great deal of the concreteness of the model of the atom proposed by Bohr and Sommerfeld. However, this lack of a definite physical picture may be the price we have to pay in order to attain the desired and necessary exactness of description of electron behavior in the atom.

The Pauli Exclusion Principle. Although four quantum numbers are necessary to specify the energy and hence the energy level of an electron in an atom, nothing said thus far would preclude several electrons from occupying the same energy level and having, therefore, all four quantum numbers the same. However, a study of spectra reveals that this is not possible, and this fact led Pauli to enunciate in 1923 his *exclusion principle*. According to this principle *no two electrons in an atom may have all four quantum numbers the same.* To exist in the same atom electrons may have three quantum numbers identical, but the fourth must be different. We shall see presently that this simple principle is of fundamental value, and aids greatly in deducing the possible electron distribution in the atoms of various elements.

Electron Shells and Subshells. In atoms containing many electrons outside the nucleus, a question of immediate interest is: How are the electrons distributed? Before turning to this question it is necessary to

develop the concept of *shells and subshells*. Just as the energy levels of the electrons in atoms may be considered to be divided into levels and sublevels, so may the electrons which occupy these be considered to exist in groups and subgroups, referred to respectively as shells and subshells. All electrons possessing the same principal quantum number n are said to be present in the same shell. Again, all the electrons in a given shell which occupy the sublevels with the same l value are said to be present in the same subshell. Thus all the electrons for which $n = 1$ and $l = 0$ occupy not only the same shell, but the same subshell. Likewise all electrons for which $n = 2$ and $l = 1$ are present in the second shell and in the $l = 1$ subshell.

This concept finds its justification in the explanation of the origin of x-ray spectra. The ordinary line spectra of elements are due to electrons present in the outermost shell, or, as they are usually referred to, the *optical electrons*. Because the optical electrons are most weakly bound to the nucleus, they are excited more readily than the electrons embedded deeper in the atom, and yield, therefore, the more easily obtainable line spectra of the elements. On the other hand, the x-ray spectra are much more difficult to excite, they involve much higher emitted energies, and hence must arise from electrons closer to the nucleus than the optical electrons. For these reasons the generally accepted theory of the origin of x-ray spectra is this. On bombardment of an atom by high-speed electrons, collisions occur between bombarding electrons and those in the atom. When the energy of the missile particles is great enough, these may dislodge electrons deep within the atom, leaving vacant spaces. To reoccupy the vacancies, electrons from higher levels will then pass into these, causing the emission of a spectrum. When this passage is from levels of $n = 2$ or greater into the $n = 1$ level, the result is the K-series of x-ray spectra. Again, when the jumps are from $n = 3$ levels or higher into the $n = 2$ level, the emission is the L-series. Similarly all transitions into the $n = 3$ level yield the M-series, those into the $n = 4$ level the N-series, etc., provided the atom contains enough electrons to yield all these series. When this is not the case, as with the lighter atoms, only the K-series, or the K and part of the L-series are observed. It will be seen that this theory of the origin of x-ray spectra ascribes their appearance not to electron jumps from sublevel to sublevel, but to electron passage from major level to major level, i.e., from shell to shell. The particular sublevel in a shell from which an electron comes determines only the fine-structure of the line, not its position. The latter is determined primarily by the change in the principal quantum number n.

With the aid of this theory, the x-ray spectra of elements can readily be accounted for, and this constitutes, therefore, a confirmation of the

concept of electron shells. By association with the series of spectra resulting from vacancies in various shells, it is customary to refer to the shell for which $n = 1$ as the K-shell, the one for which $n = 2$ as the L-shell, the one with $n = 3$ as the M-shell, etc.

The Periodic Table and Atomic Structure. If the theories of atomic structure are to be of any value to us, they must be able to explain the differences in the chemical reactivity of various elements, valence, and why the chemical and physical properties of elements repeat themselves in the manner represented by the periodic table. The periodicity of the elements definitely rules out mass as the *determining factor* in chemical reactivity, and hence chemical behavior cannot possibly be due solely to variations in the nuclei of atoms. We must seek our explanation, therefore, in the configurational architecture of the external electrons, and see what there is about it which would make the elements act as they do. It is one of the crowning achievements of modern theories of atomic structure that they have been able to shed a great deal of light on this point, and to account for the observed periodic repetition of chemical properties.

In passing from one element to another in the periodic table, the atomic number increases by one and so does the nuclear charge. To preserve the electroneutrality of the atoms this progressive increase in nuclear charge must be accompanied by a simultaneous increase in the number of orbital electrons. Consequently, in passing from hydrogen, atomic number = 1 to uranium, atomic number = 92, the number of electrons about the nucleus must increase progressively by one from a single electron for hydrogen to 92 for uranium. As these electrons enter the outer structure of the atom, they must become arranged in shells and subshells, each containing the number of electrons commensurate with the number present and the number that may be crowded into each. This much about the electronic arrangement appears to be clear. But the problems which still remain are (*a*) how many electrons may be located in each shell and subshell, and (*b*) where does each successive electron go?

The first suggestion about the number of electrons in various shells and their possible arrangement was made by Langmuir in 1919. Langmuir assumed that an electronic shell was completed every time a rare gas of the zero group in the periodic table was reached, i.e., the elements of atomic number 2, 10, 18, 36, 54, and 86. On this basis the number of electrons in the successive shells should be the differences between any two adjacent numbers, or 2, 8, 8, 18, 18, 32. However, the arrangement resulting from this distribution did not prove satisfactory. A much more fruitful proposal was advanced in 1921 independently by Bury and Bohr. These investigators proposed that the maximum number of

electrons in each shell should be given by $2\,n^2$, where n is the number of the shell, i.e., the principal quantum number. This would give for the maximum number of electrons in successive shells 2, 8, 18, 32, and 50, an arrangement quite different from Langmuir's. Further, these men suggested that *there can be present no more than eight electrons in the outermost shell of an atom before the next shell is started.* In other words, in shells where more than eight electrons can be present, only eight of these enter, then a new shell is started, and the incompleted shell is left to be filled in later. Bohr suggested also that each shell be subdivided into subshells.

That the Bury-Bohr assignment of the number of electrons to the various shells is correct can be proved with the aid of the Pauli exclusion principle and the rules given for the possible values of quantum numbers. For the K-shell we have $n = 1$, and hence the possible values which l, m, and s may have are $l = 0$, $m = 0$, and $s = +1/2$ or $s = -1/2$. Since no two electrons may have the same four quantum numbers, we arrive at the following as the only possibilities for four different quantum numbers in the K-shell:

$$n = 1 \qquad l = 0 \qquad m = 0 \qquad s = +\frac{1}{2}$$

$$n = 1 \qquad l = 0 \qquad m = 0 \qquad s = -\frac{1}{2}$$

This means that in the K-shell there is only one subshell, $l = 0$, in which only two electrons can be accommodated. On the other hand, for $n = 2$ we may have $l = 0$ and $l = 1$, $m = 0$, $m = -1$, and $m = +1$, and $s = 1/2$ or $s = -1/2$. These possibilities lead to the following combinations of four different quantum numbers:

$$n = 2 \qquad l = 0 \qquad m = 0 \qquad s = +\frac{1}{2}$$

$$n = 2 \qquad l = 0 \qquad m = 0 \qquad s = -\frac{1}{2}$$

$$n = 2 \qquad l = 1 \qquad m = 0 \qquad s = +\frac{1}{2}$$

$$n = 2 \qquad l = 1 \qquad m = 0 \qquad s = -\frac{1}{2}$$

$$n = 2 \qquad l = 1 \qquad m = -1 \qquad s = +\frac{1}{2}$$

$$n = 2 \qquad l = 1 \qquad m = -1 \qquad s = -\frac{1}{2}$$

$$n = 2 \qquad l = 1 \qquad m = 1 \qquad s = +\frac{1}{2}$$

$$n = 2 \qquad l = 1 \qquad m = 1 \qquad s = -\frac{1}{2}$$

TABLE 5 — ARRANGEMENT OF ELECTRONS IN VARIOUS ELEMENTS

At. No.	Elem.	K n=1 (0)	L n=2 (0)	L (1)	M n=3 (0)	M (1)	M (2)	N n=4 (0)	N (1)	N (2)	N (3)	O n=5 (0)	O (1)	O (2)	O (3)	P n=6 (0)	P (1)	P (2)	P (3)	Q n=7 (0)	Q (1)	Q (2)
1	H	1																				
2	He	2																				
3	Li	2	1																			
4	Be	2	2																			
5	B	2	2	1																		
6	C	2	2	2																		
7	N	2	2	3																		
8	O	2	2	4																		
9	F	2	2	5																		
10	Ne	2	2	6																		
11	Na	2	2	6	1																	
12	Mg	2	2	6	2																	
13	Al	2	2	6	2	1																
14	Si	2	2	6	2	2																
15	P	2	2	6	2	3																
16	S	2	2	6	2	4																
17	Cl	2	2	6	2	5																
18	A	2	2	6	2	6																
19	K	2	2	6	2	6		1														
20	Ca	2	2	6	2	6		2														
21	Sc	2	2	6	2	6	1	2														
22	Ti	2	2	6	2	6	2	2														
23	V	2	2	6	2	6	3	2														
24	Cr	2	2	6	2	6	5	1														
25	Mn	2	2	6	2	6	5	2														
26	Fe	2	2	6	2	6	6	2														
27	Co	2	2	6	2	6	7	2														
28	Ni	2	2	6	2	6	8	2														

Z		1s	2s	2p	3s	3p	3d	4s	4p	4d	4f	5s	5p	5d	6s
29	Cu	2	2	6	2	6	10	1							
30	Zn	2	2	6	2	6	10	2							
31	Ga	2	2	6	2	6	10	2	1						
32	Ge	2	2	6	2	6	10	2	2						
33	As	2	2	6	2	6	10	2	3						
34	Se	2	2	6	2	6	10	2	4						
35	Br	2	2	6	2	6	10	2	5						
36	Kr	2	2	6	2	6	10	2	6						
37	Rb	2	2	6	2	6	10	2	6			1			
38	Sr	2	2	6	2	6	10	2	6			2			
39	Y	2	2	6	2	6	10	2	6	1		2			
40	Zr	2	2	6	2	6	10	2	6	2		2			
41	Cb	2	2	6	2	6	10	2	6	4		1			
42	Mo	2	2	6	2	6	10	2	6	5		1			
43	Ma	2	2	6	2	6	10	2	6	5		2			
44	Ru	2	2	6	2	6	10	2	6	7		1			
45	Rh	2	2	6	2	6	10	2	6	8		1			
46	Pd	2	2	6	2	6	10	2	6	10					
47	Ag	2	2	6	2	6	10	2	6	10		1			
48	Cd	2	2	6	2	6	10	2	6	10		2			
49	In	2	2	6	2	6	10	2	6	10		2	1		
50	Sn	2	2	6	2	6	10	2	6	10		2	2		
51	Sb	2	2	6	2	6	10	2	6	10		2	3		
52	Te	2	2	6	2	6	10	2	6	10		2	4		
53	I	2	2	6	2	6	10	2	6	10		2	5		
54	Xe	2	2	6	2	6	10	2	6	10		2	6		
55	Cs	2	2	6	2	6	10	2	6	10		2	6		1
56	Ba	2	2	6	2	6	10	2	6	10		2	6		2
57	La	2	2	6	2	6	10	2	6	10		2	6	1	2
58	Ce	2	2	6	2	6	10	2	6	10	2	2	6		2
59	Pr	2	2	6	2	6	10	2	6	10	3	2	6		2
60	Nd	2	2	6	2	6	10	2	6	10	4	2	6		2
61	Il	2	2	6	2	6	10	2	6	10	5	2	6		2
62	Sm	2	2	6	2	6	10	2	6	10	6	2	6		2
63	Eu	2	2	6	2	6	10	2	6	10	7	2	6		2
64	Gd	2	2	6	2	6	10	2	6	10	7	2	6	1	2

TABLE 5 — ARRANGEMENT OF ELECTRONS IN VARIOUS ELEMENTS (Continued)

At. No.	Elem.	K $n=1$	L $n=2$		M $n=3$			N $n=4$				O $n=5$				P $n=6$				Q $n=7$		
		0	0	1	0	1	2	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2
65	Tb	2	2	6	2	6	10	2	6	10	8	2	6	1		2						
66	Dy	2	2	6	2	6	10	2	6	10	9	2	6	1		2						
67	Ho	2	2	6	2	6	10	2	6	10	10	2	6	1		2						
68	Er	2	2	6	2	6	10	2	6	10	11	2	6	1		2						
69	Tm	2	2	6	2	6	10	2	6	10	13	2	6			2						
70	Yb	2	2	6	2	6	10	2	6	10	14	2	6			2						
71	Lu	2	2	6	2	6	10	2	6	10	14	2	6	1		2						
72	Hf	2	2	6	2	6	10	2	6	10	14	2	6	2		2						
73	Ta	2	2	6	2	6	10	2	6	10	14	2	6	3		2						
74	W	2	2	6	2	6	10	2	6	10	14	2	6	4		2						
75	Re	2	2	6	2	6	10	2	6	10	14	2	6	5		2						
76	Os	2	2	6	2	6	10	2	6	10	14	2	6	6		2						
77	Ir	2	2	6	2	6	10	2	6	10	14	2	6	7		2						
78	Pt	2	2	6	2	6	10	2	6	10	14	2	6	9		1						
79	Au	2	2	6	2	6	10	2	6	10	14	2	6	10		1						
80	Hg	2	2	6	2	6	10	2	6	10	14	2	6	10		2						
81	Tl	2	2	6	2	6	10	2	6	10	14	2	6	10		2	1					
82	Pb	2	2	6	2	6	10	2	6	10	14	2	6	10		2	2					
83	Bi	2	2	6	2	6	10	2	6	10	14	2	6	10		2	3					
84	Po	2	2	6	2	6	10	2	6	10	14	2	6	10		2	4					
85	—	2	2	6	2	6	10	2	6	10	14	2	6	10		2	5					
86	Rn	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6					
87	—	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6			1		
88	Ra	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6			2		
89	Ac	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6	1		2		
90	Th	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6	2		2		
91	Pa	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6	3		2		
92	U	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6	4		2		

On this basis the L-shell should consist of eight electrons, with two of these in the $l = 0$ subshell, and six in the $l = 1$ subshell. In a like manner it can be shown that the M-shell can contain 18 electrons, with two of these in the $l = 0$ subshell, six in the $l = 1$ subshell, and ten in the $l = 2$ subshell; the N-shell 32 electrons with the subshell distribution 2, 6, 10, and 14; while the O-shell has a possible maximum of 54 electrons. These numbers are exactly those predicted by Bohr and Bury.

The Electron Arrangement in Atoms. With the aid of the Pauli exclusion principle, the Bohr-Bury postulates, and spectroscopic study, the arrangement of electrons in various atoms finally arrived at is that depicted in Table 5. In order to appreciate the significance of this table reference must first be made to the periodic classification of the elements shown in Table 6. In the simple arrangement shown in the latter table, elements are ordered according to chemical similarity in vertical *groups*, and horizontally into repetitive *periods*. The first period consists of only 2 elements, hydrogen and helium. The second and third, again, contain 8 elements each, with each period terminating with a rare gas, i.e., neon and argon. On the other hand, before the next rare gas, krypton, is reached in the fourth period 18 elements are traversed. Of these the 8 elements potassium, calcium, gallium, germanium, arsenic, selenium, bromine, and krypton behave more or less normally and exhibit fairly close similarity to preceding members of their own group, while the 10 elements starting with scandium and ending with zinc are somewhat more unusual in their behavior. The latter are generally referred to as the *transition elements*. In the fifth period the situation is exactly the same as in the fourth, the period consisting again of 18 elements, 8 of which are more or less "normal," while 10 elements, beginning with yttrium and ending with cadmium, are transition elements. However, when we come to the sixth period, we find that 32 elements are involved here. A study of these shows that they may be divided into three groups: (*a*) 8 more or less "normal" elements, cesium, barium, thallium, lead, bismuth, polonium, element 85, and radon; (*b*) 9 transition elements, lanthanum, hafnium, tantalum, tungsten, rhenium, osmium, iridium, platinum, gold; and (*c*) the rare earths listed at the bottom of the table. The latter group of elements exhibits such striking similarity in chemical properties that the only way to accommodate them in the table is to list all of them as a group in the space between barium and hafnium. Finally, the seventh period contains only the 6 remaining elements in the positions indicated.

The explanation for the periodicity of the elements and the diverse behavior of "normal," transition, and rare earth elements is embodied in the arrangement of electrons in various atoms shown in Table 5. Starting with hydrogen we see that there is only one electron present in

TABLE 6

Periodic Arrangement of the Elements

Period	Group I	Group II	Group III	Group IV	Group V	Group VI	Group VII	Group VIII	Group 0
First	1 H 1.0080								2 He 4.003
Second	3 Li 6.940	4 Be 9.02	5 B 10.82	6 C 12.010	7 N 14.008	8 O 16.0000	9 F 19.00		10 Ne 20.183
Third	11 Na 22.997	12 Mg 24.32	13 Al 26.97	14 Si 28.06	15 P 30.98	16 S 32.06	17 Cl 35.457		18 A 39.944
Fourth	19 K 39.096	20 Ca 40.08	21 Sc 45.10	22 Ti 47.90	23 V 50.95	24 Cr 52.01	25 Mn 54.93	26 Fe 55.85 27 Co 58.94 28 Ni 58.69	
Fourth	29 Cu 63.57	30 Zn 65.38	31 Ga 69.72	32 Ge 72.60	33 As 74.91	34 Se 78.96	35 Br 79.916		36 Kr 83.7
Fifth	37 Rb 85.48	38 Sr 87.63	39 Y 88.92	40 Zr 91.22	41 Cb 92.91	42 Mo 95.95	43 Ma ?	44 Ru 101.7 45 Rh 102.91 46 Pd 106.7	
Fifth	47 Ag 107.880	48 Cd 112.41	49 In 114.76	50 Sn 118.70	51 Sb 121.76	52 Te 127.61	53 I 126.92		54 Xe 131.3
Sixth	55 Cs 132.91	56 Ba 137.36	57–71* Rare Earths	72 Hf 178.6	73 Ta 180.88	74 W 183.92	75 Re 186.31	76 Os 190.2 77 Ir 193.1 78 Pt 195.23	
Sixth	79 Au 197.2	80 Hg 200.61	81 Tl 204.39	82 Pb 207.21	83 Bi 209.00	84 Po ?	85– ?		86 Rn 222
Seventh	87 — ?	88 Ra 226.05	89 Ac ?	90 Th 232.12	91 Pa 231	92 U 238.07			

* Rare Earths: 57 La (138.92), 58 Ce (140.13), 59 Pr (140.92), 60 Nd (144.27), 61 Il (?), 62 Sm (150.43), 63 Eu (152.0), 64 Gd (156.9), 65 Tb (159.2), 66 Dy (162.46), 67 Ho (164.94), 68 Er (167.2), 69 Tm (169.4), 70 Yb (173.04), 71 Lu (174.99).

the K-shell. In helium a second electron is added to this shell; and, since only two electrons can be present in any energy level with $n = 1$, this shell becomes completely occupied and closed with this element. The next element, lithium, must start, therefore, a new shell, which is the L-level, and electrons in the succeeding elements continue to occupy this shell until it also becomes completely filled when neon is reached. The only thing the added electron in sodium can do now is to start the M-shell, and the following elements up to argon continue the filling-in process in normal sequence until the first two subshells of the M-level are occupied.

But, when we come to potassium, the electron instead of filling in the $l = 2$ level of the M-shell starts a new shell. This behavior is in line with the Bohr-Bury theory that no external shell can contain more than eight electrons. Calcium follows potassium, but scandium, instead of continuing the process, starts to fill in the $l = 2$ level of the M-shell. This entrance of electrons into an inner shell continues through gallium until this subshell becomes completely occupied. Once this happens the tendency initiated by potassium is continued, and electrons again enter the outer shell in sequence until eight are present.

In the electron behavior just described lies the difference between the "normal" and the transition elements. Whereas in the "normal" elements electrons are added in sequence to the outermost electronic shell, in the transition elements the outermost shell remains essentially stationary, while succeeding electrons enter the shell immediately below the surface. Because of the filling-in process the length of the period is extended. Again, since the distribution of electrons in the outer shell is primarily responsible for the chemical properties of elements, we may anticipate that "normal" elements will act differently from the transition elements, and such is in fact the case.

The situation in the fourth period starting with rubidium and ending with xenon is essentially the same as in the third period. Like potassium, rubidium initiates a new shell, and strontium follows suit. But, starting with yttrium and continuing through cadmium, succeeding electrons enter the $l = 2$ subshell of the N-level until it is filled, then the building up of the O-shell is continued to xenon. We see here again that the transition elements arise from entrance of electrons into the shell below the surface.

The sixth period starts just like the preceding two, with cesium and barium entering the P-shell, and lanthanum the $l = 2$ sublevel of the O-shell. However, the added electron in cerium, instead of following that of lanthanum, begins to fill in the unoccupied $l = 3$ sublevel of the N-shell. Once initiated, this filling-in process of the second shell from the surface persists until it is terminated with lutecium. Hafnium con-

tinues then the process started by lanthanum, and when this is ended
with mercury, thallium and the succeeding elements proceed to build
up the *P*-shell. Reference to Table 6 will show that the elements which
so solicitously fill the gap in the *N*-shell are exactly the ones for which no
room could be found in the periodic classification, namely, the rare
earths. Since in these elements the external electronic configuration
remains unchanged from element to element, they all exhibit great
similarity of chemical properties and act as if they were a single element
in the periodic sequence.

Finally, in the seventh period a new shell is started into which enter
the electrons for element 87 and radium. With actinium begins a new
group of transition elements which terminates rapidly when uranium,
the last and heaviest element, is reached.

Thus far the emphasis has been on the reason for the periods and their
length. However, the similarity of chemical properties within groups
also follows from the table. An inspection of the *outermost* electronic
configurations of hydrogen, lithium, sodium, potassium, rubidium,
cesium, and element 87 shows that they are the same, and consist of a
single electron in a new shell. Again, the outer electronic configurations
of the elements in Group II involve two electrons, Group III three
electrons, etc. The electronic configuration of the rare gases is always
such as to necessitate in each instance the inception of a new shell by a
succeeding element. This fact indicates that a rare gas configuration
must be a highly stable arrangement of electrons, and accounts for the
reluctance these elements exhibit to enter combination with other atoms
by yielding or accepting electrons.

The Atomic Nucleus: Introduction. With this description of the
arrangement of external electrons in the various atoms, the discussion of
the external part of the atom may be considered concluded. The ques-
tion which requires attention now is the nature of the nucleus. The
only information presented thus far indicates that the nucleus consists
of protons in number equal to the atomic mass of the element, and of
electrons in number equal to the difference between the atomic mass A
and the nuclear charge Z, i.e., $(A - Z)$. No greater insight can be ob-
tained from pursuing mass spectrographic or spectroscopic studies, since
in the former the nuclei are unaffected, while in the latter only external
electrons are involved. Neither can ordinary chemical investigations be
of help, for chemical reactivity is associated again with the external elec-
trons and not the nucleus. As a matter of fact, the only phenomenon
which does involve changes in the atomic nucleus is *radioactivity*, both
natural and artificial, so we shall turn to a brief exposition of this subject.

Natural Radioactivity. In 1895 Henri Becquerel discovered that
uranium salts emit radiation which can cause fogging of photographic

plates even when the salts are wrapped in black paper. Subsequently it was shown that this radiation can ionize air, is emitted from the element as well as its salts, and is not at all affected by temperature or the source of the uranium. This spontaneous emission of radiation by an element is called *radioactivity*, while the elements which exhibit this behavior are called *radioactive elements*.

In 1898 Marie and Pierre Curie found that the mineral pitchblende (mainly U_3O_8) exhibited radioactivity more pronounced than uranium itself. This suggested the presence of elements more active than uranium, and Madame Curie actually succeeded in isolating from this mineral two new radioactive elements, polonium and radium. At about the same time thorium and actinium were also shown to be radioactive. Since that time over 30 elements have been proved to be naturally radioactive, and many others have been induced to become radioactive by artificial means to be described later.

The rays emitted by radioactive elements can be detected and measured by their ability to ionize gases. For this purpose are employed ionization chambers of the type described in Chapter IV, p. 128. Another means is to employ a Wilson cloud chamber. C. T. R. Wilson found that, when any ionizing particle passes through supersaturated water vapor, droplets of water tend to condense along the path of motion. To take advantage of this fact Wilson designed the apparatus that bears his name. This apparatus consists of a chamber filled with dust-free air and saturated with water vapor, and fitted with a piston. As the particles enter through a window the air is made to expand, the temperature drops, and the vapor becomes supersaturated. As a result a fog track deposits along the path of the particles which can be photographed. From such fog tracks the velocity and energy of the particles can be ascertained.

Nature of Radioactive Emission. The rays emitted by radioactive elements and their salts have been shown to consist of three types of radiation, α, β, and γ rays. Study of the α rays in a magnetic field has revealed that these consist of material particles possessing a mass four times and a *positive charge* twice that of the proton. This would make the charge and mass of the α particles identical with that of the *helium nucleus*, i.e., doubly charged helium atom, and this identity has actually been established. As we have seen, these particles have the ability to ionize gases and can penetrate matter, as Rutherford showed in his scattering experiments. The initial velocity of the α particles on emission is very high, ranging from $1-2 \times 10^9$ cm per second as against the velocity of light 3×10^{10} cm per second.

Like α rays, β rays have been shown to consist of particles. However, the β rays have a much lower mass and bear a negative charge. Their

e/m ratio is identical with that of the electron, and this, in fact, is what they are, *electrons*. These rays again can ionize gases, affect a photographic plate, and penetrate matter. But whereas their penetrating power is about 100 times that of α particles, their ionizing power is only about 1/100 as great because of their much lower mass. Finally, the velocity of the emitted electrons depends on their source and is very large, approaching in some instances very close to the velocity of light.

γ rays are quite different from either the α or β rays. They are totally unaffected by electric and magnetic fields, and behave in every respect as electromagnetic radiation of the same nature as x rays. Measurements of their wave length indicate that they are of the order of $10^{-8} - 10^{-11}$ cm in length, and hence are shorter than x radiation. Like the other rays, they affect photographic plates, ionize gases, and penetrate matter. In penetrating power they are about 10 to 100 times as effective as β rays, but in ionizing power they are proportionately weaker.

Theory of Radioactive Disintegration. The fact that radioactive elements emit α particles indicates that the portion of the atom involved in the emission cannot possibly be the external part, but must be the *nucleus*. To account for this nuclear radiation Rutherford and Soddy advanced in 1903 their theory of *radioactive disintegration*. Rutherford and Soddy proposed that the nuclei of radioactive elements, unlike those of the stable elements, are unstable, and decompose spontaneously by emission of an α or β particle to form a *new element* of different chemical and physical properties. Thus, when uranium-I with $A = 238$ and $Z = 92$ emits an α particle to yield uranium-X_1, the mass should decrease by four units to $A = 234$, and the nuclear charge by two units to $Z = 90$. Since the nuclear charge has decreased, two external electrons which are no longer necessary will also be lost, and we obtain therefore from the disintegration a new element of $Z = 90$ and $A \doteq 234$, which is, evidently, an isotope of thorium. Similarly when uranium-X_1 emits a β particle, the nuclear mass remains unchanged, but the nuclear charge is increased to $Z = 91$. The new product, uranium-X_2, has now a mass of 234 and an atomic number of 91, and hence is an isotope of protoactinium. This idea that as a result of radioactive spontaneous disintegration new elements are formed has been amply verified, and is a well-established working hypothesis in the field of radioactivity.

From the examples cited it may be observed that the product of a radioactive decay may itself be radioactive and undergo further decomposition. This, as a matter of fact, is generally the case. For example, uranium-I yields a whole series of consecutive disintegration products until eventually radium-G, an isotope of lead, is reached which is stable. Analogously, thorium goes through a sequence of disintegrations until the stable thorium-D, again an isotope of lead, is attained. Each of the

intermediate steps involves the emission of either an α or a β particle, and only rarely both. The γ radiation arises from energy rearrangements within the nucleus after an α or β particle ejection. When the rearrangement leads to a lower stable potential energy, the excess is evolved as a quantum of energy of the frequency of the γ ray. Not all nuclear transformations are accompanied by emission of γ rays.

The transformation brought about by the emission of an α or β particle from the nucleus is summarized in the Fajans-Soddy-Russell *displacement law*. This law states that whenever a parent nucleus emits an α particle its atomic number is decreased by *two units*, and the new element is shifted two positions *to the left* in the periodic table from that of the parent. On the other hand, when the parent nucleus emits a β particle, the atomic number is increased by *one*, and hence the product is shifted one place *to the right* of its parent in the periodic table. This law is merely an epitome of the changes described in the examples cited in the first paragraph of this section.

Rate of Radioactive Decay. All radioactive disintegrations proceed at rates which cannot be modified by any chemical or physical device known to man. For any given element this rate at any instant of time is proportional only to the number N of the nuclei of the element present, namely,

$$-\frac{dN}{dt} = kN \tag{26}$$

where k is a rate constant characteristic of the element. Equation (26) is identical with that for a first order chemical rate expression. Its integration between the limits $N = N_0$ at $t = 0$ and $N = N$ at $t = t$ yields

$$t = \frac{1}{k} \ln \frac{N_0}{N} \tag{27}$$

from which the half-life period, $t_{1/2}$, of the radioactive element follows as

$$t_{1/2} = \frac{1}{k} \ln \frac{N_0}{\dfrac{N_0}{2}}$$

$$= \frac{\ln 2}{k} \tag{28}$$

Equation (28) shows that the time necessary for one-half an initial quantity of radioactive substance to be disintegrated is a constant independent of the initial amount present, and is characteristic only of the element in question. For the various naturally radioactive species $t_{1/2}$ may vary from 10^{-11} sec to 1.3×10^{10} years!

Radioactive Series. Naturally radioactive elements belong to one of three radioactive series in which the parent elements are either uranium, thorium, or protoactinium, and in which the end products of final disintegration are all isotopes of lead. These series bear in each instance the name of the parent element. Many of the radioactive elements in any given series are isotopes of each other, as well as of members of other series.

The Neutron. Although α particles and electrons are the only material corpuscles emitted by naturally radioactive substances, these are not the only particles whose presence in an atomic nucleus is indicated. In considering the cause of nuclear stability on the basis of existence of protons and electrons in the nucleus, Rutherford in 1920 came to the conclusion that there may be present in the nucleus also a particle of the *same mass as the proton but with no charge.* The first evidence of the actual existence of such uncharged particles, called *neutrons,* did not come forth until 1930. In that year Bothe and Becker observed that when light atoms such as lithium and beryllium are bombarded with α particles, a very penetrating radiation is emitted. Immediately thereafter Irene Curie Joliot, daughter of Madame Curie, and F. Joliot, her husband, showed that this radiation has the power to release protons from materials containing a large number of hydrogen atoms, such as paraffins. This led Chadwick to believe that neutrons may be present in the emitted radiation, and to establish in 1932 that such is the case.

Chadwick's conclusion has been confirmed by cloud chamber studies and by the fact that the penetrating radiation under discussion is unaffected by electric and magnetic fields, indicating absence of charge. It has also been shown that the mass of the neutron is 1.00894, which is almost identical with that of the positively charged proton, 1.00813. As neutrons bear no electric charge, they are neither attracted nor repelled by electrons or the positively charged nuclei, and so they can pass through large thicknesses of various materials, such as lead, without appreciable stoppage or reflection. Again, because they are uncharged, neutrons can approach charged nuclei very closely, and by severe collision can dislodge protons from light atoms such as hydrogen by transfer of momentum.

Chadwick generated his neutrons by bombarding a beryllium target with α particles from radioactive polonium. The production of neutrons by bombardment of beryllium with α particles is due to the *nuclear reaction.*

$$^9\text{Be} + {}^4\text{He} \longrightarrow {}^{12}\text{C} + {}^1n \tag{29}$$
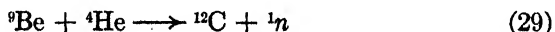
The superscripts to the left of elements designate the nuclear mass of each specie involved, while n stands for the neutron, whose mass is

essentially unity. The reaction given involves the capture of an α particle by the beryllium nucleus with the resultant formation of carbon and the emission of a neutron. The process is, therefore, a *transmutation of elements*, the dream and goal of medieval alchemists. Other important methods of neutron generation are the bombardment of deuterium atoms with deuterons, i.e., ionized deuterium atoms, in which instance the products are ^{3}He, an isotope of helium, and a neutron,

$$^2D + \,^2D \longrightarrow \,^3He + \,^1n \tag{30}$$
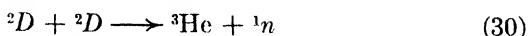
the bombardment of lithium with deuterons to form beryllium and a neutron,

$$^7Li + \,^2D \longrightarrow \,^8Be + \,^1n \tag{31}$$
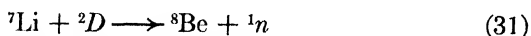
and the interaction of deuterium with γ radiation to produce hydrogen and a neutron, namely,

$$^2D + \gamma \text{ rays} \longrightarrow \,^1H + \,^1n \tag{32}$$
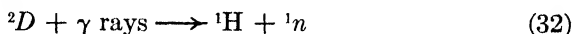
The Positron. The occurrence of a unit of negative electricity with a very low mass, i.e., the electron, would suggest that there may be present in matter also a corresponding positive particle of low mass, namely, a *positron*. However, even though Dirac showed theoretically in 1928 that positrons should exist, all attempts to find such a particle proved fruitless until 1932. In that year, while studying cosmic rays with the aid of a Wilson cloud chamber, Carl Anderson detected in the cloud chamber a fog track which could only be produced by a positively charged particle of very low mass. Anderson ascribed this track to the passage of a positron. This conclusion was confirmed quickly by other workers, and the existence of positrons was thus definitely established.

The positron has been shown to have a mass identical with that of the electron, and to bear a charge equal in magnitude but opposite in sign. However, whereas the electron is stable, the positron is very short-lived. When released from a nucleus a positron loses its kinetic energy very rapidly, in 10^{-10} sec or so, and combines then with an electron to cause the annihilation of both particles. To compensate for the destroyed mass 2 quanta of x radiation are formed whose total energy is equal to the mass destroyed in accordance with the Einstein relation $E = mc^2$.

Various methods are available at present for the production of positrons. When beryllium, magnesium, or aluminum are bombarded by α particles, these elements become *artificially radioactive*, and emit positrons. In each instance the process involves two consecutive nuclear reactions, the first of which is the formation of an artificially radioactive element, the second the emission of the positron. For aluminum these reactions are:

$$^{27}_{13}Al + \,^4_2He \longrightarrow \,^{30}_{15}P + \,^1_0n \tag{33a}$$

$$^{30}_{15}P \longrightarrow \,^{30}_{14}Si + e^+ \tag{33b}$$
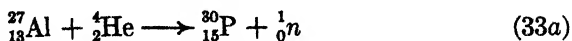
In these equations the superscripts give again the nuclear masses, the subscripts the nuclear charges. e^+ is the symbol for the positron, i.e., a positive electron. Thus equation (33a) states that aluminum, with mass = 27 and nuclear charge = 13, reacts with an α particle, mass = 4 and atomic number = 2, to form phosphorus, of mass = 30 and charge = 15, and a neutron of unit mass and zero charge. Again equation (33b) states that the phosphorus is radioactive and decomposes to yield a positron and silicon of mass = 30 and atomic number = 14.

Similar reactions result also from bombardment of nitrogen, fluorine, sodium, phosphorus, and potassium with α particles, from the bombardment of various elements with protons or deuterons, and from the interaction of lead with γ rays.

Structure of the Nucleus. The emission of α particles by radioactive nuclei and the discovery of neutrons and positrons calls for a revision of our concept of the structure of the nucleus. Whereas in the early days of atomic structure research cognizance had to be taken only of possible presence of protons and electrons in the nucleus, now not only these but also neutrons, positrons, and α particles must be considered. At present the belief is that atomic nuclei are composed of neutrons and protons. In hydrogen only a single proton is present. In deuterium, on the other hand, the nucleus is thought to consist of a neutron and a proton, while in helium, i.e., the α particle, the composition is two neutrons and two protons in intimate combination. In the same way may be built up the more complex nuclei. Just where the positron and electron come in is not so certain. Indications are that the latter two particles are not present in the nucleus as such, but result from dissociations of protons and neutrons. If we imagine the proton to consist of a neutral particle and a positron, and the neutron of a neutral particle plus an electron and positron, the situation in the nucleus may be understandable.

Transmutation of Elements. In the preceding discussion various examples were given of the transmutation of one element into another. Until about 1919 the only instances of such transformations were the natural radioactive disintegrations in which a more complex element changed into simpler elements by emission of either α or β particles. All these processes were spontaneous, and could not be controlled in any known manner. However, since 1919 various means have been developed by which transmutations of elements can be brought about *artificially*. These all entail the bombardment of elements with relatively heavy and high-speed atomic particles such as α rays, neutrons, protons, and deuterons as the transmuting projectiles. For certain processes high-energy γ rays are also suitable. Until relatively recently the only sources of such high-speed, and therefore high-energy, missiles were

the α particles emitted by naturally radioactive elements. This is no longer true. At present apparatus is available, such as powerful transformers, electrostatic generators, high voltage discharge tubes, and the cyclotron developed by Lawrence at the University of California, by means of which charged particles can be formed and accelerated to terrific velocities. Thus far Lawrence has been able to construct cyclotrons that yield energies equivalent to accelerating electrons through an electrostatic field of 16,000,000 volts. At present a huge cyclotron is under construction at the University of California by which, it is expected, fields equivalent to 100 to 300 million volts may be attained. Such a cyclotron should yield deuteron beams of sufficient penetrating power to pass through 140 feet of air as against 5 feet now.

Hundreds of transmutations have been accomplished by such bombardments. When the bombarding missiles are α particles, a nuclear transformation takes place which results in the emission of neutrons or protons. With protons as projectiles the products may be α particles, neutrons, deuterons, or γ rays and a new element. Neutron bombardment, in turn, may release α particles, protons, two neutrons, or γ rays; deuteron bombardment α particles, protons, neutrons, and a combination of α particles and protons or α particles and neutrons; while γ rays usually yield neutrons.

Artificial Radioactivity. In 1934, I. and F. Joliot discovered that some products of nuclear reactions are unstable and tend to change to more stable nuclei by *spontaneous* positron emission. Since that time over 350 such elements have been observed, and it has been found that the emissions may involve not only positrons, but also electrons and γ radiation. Since this type of radioactivity does not arise spontaneously but must be brought about by preliminary appropriate bombardment, the phenomenon is referred to as *artificial or induced radioactivity*. In each instance the elements exhibiting induced radioactivity are unstable isotopes of the known elements. They do not appear among the stable isotopes because their life period is usually short, and hence it is hardly to be expected that they will be found in nature.

Once produced, artificially radioactive elements behave very much like those which are naturally radioactive. They disintegrate in each instance according to the same rate law, equation (26), and exhibit thus constant half-life periods characteristic only of the element in question. Further, although some artificially radioactive elements may be produced by several different methods, once formed, each element decomposes at the same rate and has thus the same life period. There is, however, one respect in which the two do differ. Whereas natural radioactivity is confined only to the very massive nuclei, artificial radioactivity can be induced in light nuclei as well. Thus not only ^{210}Bi,

^{209}Pb, ^{205}Pb, ^{206}Tl, and ^{198}Au are artificially radioactive, but also such light nuclei as ^{3}H, ^{6}He, ^{7}Be, and ^{10}C.

In Table 7 are given some examples of artificially radioactive elements, the methods by which they are produced, their half-life periods, and the nature of the emitted radiation.

TABLE 7

SOME ARTIFICIALLY RADIOACTIVE ELEMENTS

Element	Method of Production	Half-Life Period	Emitted Radiation *
$^3_1$H	$^2_1D + {}^2_1D$	> 10 years	e
$^6_2$He	$^9_4\text{Be} + {}^1_0n$	0.8 sec	e
$^{12}_5$B	$^{11}_5\text{B} + {}^2_1D$	0.022 sec	e
$^{11}_6$C	$^{10}_5\text{B} + {}^2_1D$	20.5 min	e^+
$^{14}_6$C	$^{13}_6\text{C} + {}^2_1D$	$\gg 1000$ years	e
$^{13}_7$N	$^{12}_6\text{C} + {}^1_1\text{H}$	9.93 min	e^+
$^{15}_8$O	$^{14}_7\text{N} + {}^2_1D$	2.21 min	e^+
$^{24}_{11}$Na	$^{23}_{11}\text{Na} + {}^2_1D$	14.8 hours	e
$^{30}_{15}$P	$^{27}_{13}\text{Al} + {}^4_2\text{He}$	2.55 min	e^+
$^{41}_{18}$A	$^{40}_{18}\text{A} + {}^1_0n$	110 min	e
$^{38}_{19}$K	$^{39}_{19}\text{K} + {}^1_0n$	7.7 min	e^+
$^{80}_{35}$Br	$^{79}_{35}\text{Br} + {}^1_0n$	18 min	e^+
$^{124}_{53}$I	$^{121}_{51}\text{Sb} + {}^4_2\text{He}$	4.0 days	e^+
$^{198}_{79}$Au	$^{197}_{79}\text{Au} + {}^1_0n$	2.7 days	e

* e = electron; e^+ = positron.

Nuclear Fission. For the nuclear reactions described above the change in mass produced by bombardment of a nucleus is always small. In each case the nucleus either increases in mass by capture of a particle, or loses some mass by emission of a particle, but never does it rupture to yield fragments of appreciably smaller masses. However, in 1934 Enrico Fermi discovered that uranium when bombarded with neutrons yielded electrons and other products of decay. In 1939 Hahn and Strassmann showed that what happens here is a *rupture* of the uranium nucleus to form fragments which have a mass and atomic number considerably lower than that of uranium. Thus they showed that two of the major fragments resulting from uranium (atomic number = 92) fission were barium (at. no. = 56) and lanthanum (at. no. = 57). This process is, therefore, not a relatively minor nuclear transformation, but a drastic neutron bombardment initiated smashing of a nucleus into smaller fragments.

Further research has established that of the various uranium isotopes the one primarily responsible for the fission is ^{235}U. Accompanying the

breakup of this nucleus is the emission of a tremendous amount of atomic energy which is associated with the various fragments. Measurements indicate that the total energy thus released may be as high as 175 million electron-volts, or 4×10^{12} calories per gram atom ruptured! The liberation of such high energies immediately raises the prospect of the utilization of atomic energy as a source of power, a long and cherished hope. But the realization of this hope is faced with a difficulty. Ordinary uranium consists of 99.28 per cent ^{238}U, 0.71 per cent ^{235}U, and about 0.01 per cent ^{234}U. To obtain from this isotopic mixture the desired ^{235}U, a method of separation is required which can yield appreciable quantities of this element. Although some methods of separation, such as the mass spectrograph, have been used and although a great deal of research is still in progress, no practical scheme of ^{235}U concentration has as yet been announced. Until such a method is developed and sizeable amounts of ^{235}U become available, the dream of utilizing atomic energy as a source of power may have to remain a dream. Nevertheless, the prospect is an intriguing one, and there is, of course, the possibility that some alternate means of tapping atomic energy may be discovered.

Applications. Many of the isotopes and nuclear particles described in this chapter find extensive application in the fields of physics, chemistry, biology, and medicine. We have already seen how α particles, protons, neutrons, and deuterons can serve as projectiles for the transmutation of elements, for the production of artificial radioactivity, and for nuclear fission. γ rays are employed in medicine for the treatment of cancer as well as in nuclear studies. In the former application radon resulting from radium decomposition is collected, compressed, and sealed into tiny ampoules which are then used to treat the infected areas.

Radioactive elements, both natural and artificial, have been employed to determine the·solubility of difficultly soluble salts and to estimate the extent of and the surface area involved in adsorption. Both of these applications depend on the fact that a radioactive element added to the same stable element will persist in a definite ratio irrespective of the state the element may be in. Consequently, when, say, a trace of radioactive lead is added to a solution of ordinary lead, and the lead is precipitated as either the sulfate or chromate, a determination of the radioactivity of the filtrate readily yields the concentration of lead still dissolved. Similarly from the activity of a material adsorbed from a mixture of stable and radioactive isotopes it is possible to estimate the extent of adsorption, the number of adsorbed molecules, and the area covered by these.

Isotopes, stable and radioactive, have been used to trace the course of chemical and physiological reactions. Used in this manner these substances are called *isotopic tracers*. Deuterium, and the isotopes of carbon,

oxygen, and nitrogen, have been used to elucidate the mechanism of kinetic reactions, and to follow the course these elements take in passing through an animal body. In the first instance the rates of reactions with the ordinary isotopic species are measured, then the heavier or lighter isotopes are incorporated, and the rates are measured again. From such data it is frequently possible to infer the path followed by the reaction, and hence the mechanism. In physiologic work the uncommon isotopes are incorporated in various foods, as D and O in water or the carbohydrates, N in proteins, C in carbohydrates or fats, fed to animals, and the concentration of these in various organs estimated. From the accumulation of any of these tracers in an organ it is possible to decide just where each food goes. For these purposes radioactive isotopes are also of value, for they permit the tracing of a food by the emission intensity. In this manner have been used radioactive phosphorus and iodine.

Another interesting application of isotopes is in *exchange reactions*. When a compound containing hydroxyl or carboxyl groups is placed in heavy water, it is found that deuterium exchanges itself for ordinary hydrogen atoms in these compounds. This fact indicates that the hydrogen atoms in these compounds are not static, but ionize to be replaced by the more plentiful deuterium atoms. Similar exchange has been found in the hydrogens of acetylene, acetone, and chloroform, but not in some of the paraffin hydrocarbons. In like manner various halogens have been found to exchange with their ions in solution, indicating a mobile equilibrium in action, but not with alkyl halides. A particularly interesting fact revealed by exchange studies is that sulfur atoms in sulfates do not exchange while those in thiosulfates do, and that whereas no exchange can take place between manganous and permanganate ions, interchange does occur rapidly between manganous and manganic ions.

REFERENCES FOR FURTHER READING

1. F. W. Aston, "Mass Spectra and Isotopes," Edward Arnold & Co., London, 1933.
2. S. Dushman, "The Elements of Quantum Mechanics," John Wiley & Sons, Inc., New York, 1938.
3. A. Farkas, "Ortho-, Para-, and Heavy Hydrogen," Cambridge University Press, Cambridge, Eng., 1935.
4. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.
5. G. Herzberg, "Atomic Spectra and Atomic Structure," Prentice-Hall, Inc., New York, 1937.
6. G. von Hevesy, "Applications of Radioactive Indicators to Biology," Ann. Rev. Biochem., **9**, 641 (1940).
7. R. A. Millikan, "Electrons (+ and −), Photons, Neutrons, and Cosmic Rays," University of Chicago Press, Chicago, 1935.

8. Pauling and Wilson, "Introduction to Quantum Mechanics," McGraw-Hill Book Company, Inc., New York, 1935.
9. Physics Staff, University of Pittsburgh, "Outline of Atomic Physics," John Wiley & Sons, Inc., New York, 1937.
10. Rutherford, Chadwick, and Ellis, "Radiations from Radioactive Substances," Cambridge University Press, Cambridge, Eng., 1930.
11. G. T. Seaborg, "Artificial Radioactivity," Chem. Rev., **27**, 199 (1940).
12. L. A. Turner, "Nuclear Fission," Rev. Mod. Physics, **12**, 1 (1940).
13. Taylor and Glasstone, "Treatise on Physical Chemistry," third edition, D. Van Nostrand Company, Inc., New York, 1942, Vol. I.

PROBLEMS

1. One volume of hydrogen gas combines with one volume of chlorine gas to give two volumes of hydrogen chloride gas. Show that the assumption that hydrogen and chlorine are diatomic is in accord with the observed facts, while the assumption that these gases are monatomic or triatomic would be contradictory.

2. A water droplet of unit density and radius 10^{-4} cm captures three unit charges of electricity. What electric field in volts/centimeter will be required to keep the drop at rest? One electrostatic unit of potential is equal to 300 volts. If the distance between the parallel charged plates is 2.40 cm, what must be the potential difference across the plates?   *Ans.* 855 volts/cm; 2052 volts

3. Using Stokes's law find the mass and radius of a water droplet which falls freely in air with a velocity of 0.5 cm/second. The viscosity coefficient of air is 0.00018 cgs units.   *Ans.* 6.43×10^{-4} cm, 1.11×10^{-9} g

4. Calculate the mass in grams of an NH_3 molecule.

5. What will be the mass of an electron moving with (*a*) a velocity of 5×10^9 cm/second? (*b*) 2.9×10^{10} cm/second?

6. Ascertain whether there would be an increase or decrease in energy accompanying the transformation of two atoms of H to an atom of D. What would be the magnitude of the change in kilocalories?

7. Calculate the amount of energy in calories which would be emitted per second at 2000° K from a perfect black body 1 sq mm in area.   *Ans.* 0.434 cal

8. Calculate the frequency in wave numbers and the wave length in angstrom units of (*a*) the first two lines of the Lyman series, (*b*) the third and fourth lines of the Paschen series.

9. The wave length of a certain line in the Balmer series is observed to be 4341 Å. To what value of n_2 does this correspond?   *Ans.* 5

10. (*a*) What total amount of energy in calories would be required to shift all the electrons from the first Bohr orbit to the sixth Bohr orbit in a gram atom of hydrogen atoms? (*b*) Through what distance would each electron have to move? (*c*) What frequency of radiation would be emitted if the electrons returned to their initial state?

11. What is the frequency and wave length of radiation emitted by the transition of the electron from the second to the first Bohr orbit in singly ionized helium? How do these compare with the corresponding values for the hydrogen atom? Assume that R is the same as the Rydberg constant R_H for hydrogen.

12. What possible values may the other quantum numbers assume when the principal quantum number is 5? According to the Pauli principle, what maximum number of electrons may be found in each of the l subshells?

13. Explain the fundamental difference between the electronic configurations of the Group I *A* elements (Li, Na, K, etc.) and those of the Group I *B* elements (Cu, Ag, Au).

14. One gram of Ra is placed in an evacuated tube whose volume is 5.0 cc. Assuming that each Ra nucleus yields 4 He atoms which are retained in the tube, what will be the partial pressure at 27° C of the He produced at the end of a year? For Ra the period of half-life is 1590 years.   *Ans.* 0.96 mm Hg

15. Assuming that the nucleus is composed of protons and neutrons, what would be the number of each in Pb-206? What would be the number of α particles and neutrons in this nucleus if it were made up of such?

16. Aluminum when bombarded with α particles emits protons. Write a balanced equation for the nuclear reaction involved. Write an equation for the production of neutrons from B by bombardment with α particles.

17. In what period of time would a sample of $^{24}_{11}Na$ lose 90% of its intensity?

Ans. 49.2 hr

Molecular Structure

Valence. Atoms in reacting with each other to form molecules vary greatly in their combining power. To express these various reactivities the concept of *valence* has been introduced, which expresses in terms of whole numbers the relative reacting powers of the various elements. In this scheme the common valences of hydrogen and the alkali metals are $+ 1$, those of the alkaline earth metals $+ 2$, of aluminum and boron $+ 3$, carbon and silicon $+ 4$, etc. On the other hand, the valences of the elements to the right of the periodic table, i.e., of the nonmetals, are negative, and are $- 1$ for the halogens, $- 2$ for oxygen and sulfur, etc.

The *periodic* recurrence of chemical properties of the elements on increase of atomic number, and hence mass, rules out the nucleus and points to the extranuclear electrons as the seat of chemical reactivity. In discussing the distribution of electrons about various nuclei it was shown that, independently of the total number of electrons present the electronic configuration in the external shells of all atoms belonging to a given period was essentially identical. The latter fact strongly suggests that chemical reactivity is associated primarily with the outermost electrons in an atom, and does not depend to any degree on electrons in shells much below the surface. Any explanation for the laws of chemical combination and the valences of various elements must be sought, therefore, in the behavior of the outermost electrons, and in their interaction with the similar electrons of other atoms.

The Lewis-Kossel Theory of Valence. The first attempt to explain valence in terms of electrons was made by W. Kossel in 1916. The starting point of Kossel's theory is the observation that the great stability of the inert gases is attained when the external shell of the atom contains two electrons in helium and eight electrons in every other case. Kossel suggested, therefore, that all atoms tend to reach rare gas configurations either by taking on or losing electrons. For example, an atom of sodium contains two closed shells with one electron outside of these. Disregarding the closed shells and writing a dot to represent an electron, the neutral sodium atom may be designated by Na. Again,

chlorine contains two closed shells of electrons with seven electrons outside of these, namely, $: \overset{\cdot\cdot}{\underset{\cdot\cdot}{Cl}} \cdot$. Now, to attain a rare gas configuration sodium can either divest itself of the one external electron it possesses and become the positive ion Na^+, or it may acquire seven electrons and become Na^{-7}. Obviously the first of these processes should be easier to accomplish, and so the tendency of sodium will be to lose an electron and become Na^+. On the other hand the chlorine atom may acquire a rare gas configuration either by gaining one electron or by losing all seven. Here evidently the first process is the more probable, and hence chlorine would tend to form the ion Cl^-. These natural tendencies come into play when sodium and chlorine are brought together. We obtain thus the reaction

$$\overset{\cdot}{Na} + \cdot \overset{\cdot\cdot}{\underset{\cdot\cdot}{Cl}} : \; = Na^+ + \left[: \overset{\cdot\cdot}{\underset{\cdot\cdot}{Cl}} :\right]^-$$
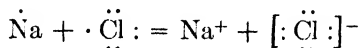
in which an electron is transferred from the sodium to the chlorine to yield a sodium and a chlorine ion, each with a rare gas configuration. Similarly the reaction of barium and sulfur can be represented by the electron transfers

$$\overset{\cdot}{Ba} + \overset{\cdot\cdot}{\underset{\cdot}{S}} : \; = Ba^{++} + \left[: \overset{\cdot\cdot}{\underset{\cdot\cdot}{S}} :\right]^{--}$$
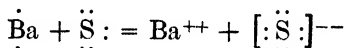
This type of atomic interaction involving the *outright transfer* of one or more electrons from one atom to another leads to the formation of ions which are held together by electrostatic attraction. Because of the electrostatic nature of the binding force, the bond between the atoms is said to be *electrovalent*, and the valence exhibited is said to be an *electrovalence*. This is the type of binding found in simple salts such as sodium chloride, potassium iodide, ammonium chloride, and barium nitrate. However, although this explanation is satisfactory for such substances, it can hardly account for the formation of molecules such as CH_4, NO, H_2, N_2, and others of this category. In these no ions can be detected, and no reason is immediately apparent why, for instance, two atoms of hydrogen should combine into a molecule. To get around this difficulty G. N. Lewis proposed in 1916 that not only can union between atoms arise from transfer of electrons, but a rare gas configuration can be attained also during reaction from *sharing electrons in pairs*. For example, if one hydrogen atom with a single external electron were to share this electron with another hydrogen atom, the result would be a hydrogen molecule

$$H \cdot + \cdot H = H : H$$

in which each hydrogen atom could claim the pairs and possess thereby the helium structure. Again, the formation of a molecule such as CH_4

would arise from the sharing of electrons between hydrogen and carbon as follows

$$4\,\text{H} \cdot + \cdot \overset{\cdot}{\underset{\cdot}{\text{C}}} \cdot = \text{H} : \overset{\text{H}}{\underset{\text{H}}{\overset{..}{\text{C}}}} : \text{H}$$
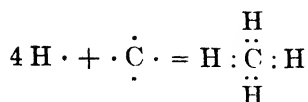
Here again each hydrogen attains by sharing electrons the rare gas configuration of helium, while the carbon by being able to share in its own as well as the four acquired electrons has the rare gas configuration of neon. In all such sharings *every electron pair corresponds to a single valence bond*. To differentiate the sharing type of atomic combination from electrovalent union, the terms *covalence* and *covalent bond* are employed. In this connection it is of interest to point out that with very few exceptions molecules contain an even number of electrons.

Covalent binding resulting from electron sharing can be used also to account for the formation of double and triple valence bonds in molecules. Thus the formation of an oxygen molecule can be represented by

$$\overset{..}{\text{O}} : + : \overset{..}{\underset{..}{\text{O}}} = \overset{..}{\text{O}} : : \overset{..}{\underset{..}{\text{O}}}$$
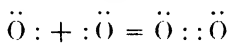
in which instance four electrons, corresponding to two valence bonds, are shared by the two atoms. Similarly the electronic structure of ethylene can be given as

$$\text{H} : \underset{\overset{..}{\text{H}}}{\text{C}} : : \underset{\overset{..}{\text{H}}}{\text{C}} : \text{H}$$
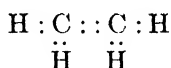
where the four electrons between the two carbons indicate the doubled bond. Finally, in triple bonds three electron pairs are shared as in N_2, $: \text{N} : : : \text{N} :$, or in C_2H_2, $\text{H} : \text{C} : : : \text{C} : \text{H}$. In the one instance each nitrogen contributes three electrons to the formation of the covalent triple bond, while in the other each carbon supplies three electrons.

A covalent bond can arise also from the sharing of electrons *supplied by one atom only*. The bond then is said to be *coordinate covalent*, or that we have *coordinate covalence*. The atom contributing the electron pair to form the bond is further called the *donor atom*, the one to which the pair is donated the *acceptor atom*. An example of this is found in the formation of the NH_4^+ ion from ammonia and the hydrogen ion, namely,

$$\text{H} : \overset{\text{H}}{\underset{\overset{..}{\text{H}}}{\overset{..}{\text{N}}}} : + \text{H}^+ = \left[\text{H} : \overset{\text{H}}{\underset{\overset{..}{\text{H}}}{\overset{..}{\text{N}}}} : \text{H} \right]^+$$
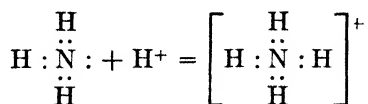
In the ammonia molecule four electron pairs are present, of which only three are shared. When a hydrogen ion comes near this molecule the nitrogen allows the proton to share with it the free pair of electrons,

and the result is thus the formation of the ammonium ion. Once established by coordinate covalent sharing of this kind, the bond between the acceptor hydrogen and the nitrogen is apparently indistinguishable from the other N—H bonds, and the positive charge becomes the property of the whole group rather than of any one hydrogen. As other examples of coordinate covalent binding may be cited the formation of the complex $BF_3 \cdot NH_3$ from BF_3 and NH_3,

$$
\begin{array}{ccc}
:\overset{..}{F}: & H & :\overset{..}{F}:H \\
:\overset{..}{F}:\overset{..}{B} \; + \; :\overset{..}{N}:H \; = \; :\overset{..}{F}:\overset{..}{B}:\overset{..}{N}:H \\
:\overset{..}{F}: & H & :\overset{..}{F}:H
\end{array}
$$
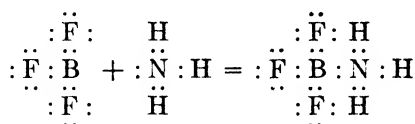
and of BF_4^- from BF_3^- and F^-, namely,

$$
\begin{array}{cc}
:\overset{..}{F}: & :\overset{..}{F}: \\
:\overset{..}{F}:\overset{..}{B} \; + \; :\overset{..}{F}:^- \; = \; \left[:\overset{..}{F}:\overset{..}{B}:\overset{..}{F}: \right]^- \\
:\overset{..}{F}: & :\overset{..}{F}:
\end{array}
$$
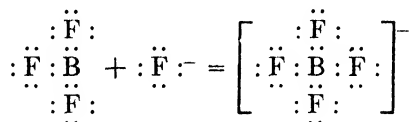
In both these instances the acceptor atom, boron, possesses only six electrons about it. In order to reach the stable octet of electrons it accepts an electron pair from the donors NH_3 and F^- to form the respective complexes. By this means both the donor and acceptor atoms attain rare gas configurations which are conducive to maximum all-around stability.

Nature of Bond and Physical Properties. The manner in which a chemical bond arises reflects significantly in the properties of the resulting compounds. Because electrovalency leads to the formation of ions held together by electrostatic attraction, compounds resulting from this type of binding are ionized in the solid as well as in the molten states. In the latter state they are good conductors of electricity; and on solution in such solvents as water they yield strong electrolytes which also conduct a current. Further, because of the strong electrostatic forces between the electrically charged ions the compounds exhibit relatively high melting and boiling points. In contrast to the electrovalent compounds those resulting from covalency are nonionized and hence non-conducting, and they exhibit much lower melting and boiling points. Again, whereas electrovalent compounds are generally more soluble in polar solvents such as water or ethyl alcohol, covalent compounds are much more readily soluble in typical organic solvents such as benzene, ether, or carbon tetrachloride. These differences in the two types of binding may be judged from the quantitative data given in Table 1 for several electrovalent and covalent compounds.

TABLE 1

COMPARISON OF PROPERTIES OF ELECTROVALENT AND COVALENT COMPOUNDS

Property	Electrovalent		Covalent	
	NaCl	BaCl$_2$	CCl$_4$	CH$_4$
Melting point	800° C	962° C	− 22.6° C	− 182.6° C
Boiling point	1413° C	1560° C	76.8° C	− 161.4° C
Conductance at m.p.	133	64.6	0	0
Solubility in H$_2$O at 20° C (g/100 g H$_2$O)	36.0	35.7	Insoluble	Insoluble
Solubility in ether	Insoluble	Insoluble	∞	104 cc at 10° C

The products resulting from coordinate covalent binding may, as we have seen, be both uncharged molecules, as BF$_3$ · NH$_3$, or ions, as NH$_4^+$ or BF$_4^-$. In this category belong also the various other coordinated compounds, such as the hydrates, ammoniates, alcoholates, etherates, etc. For this variety of compounds it is very difficult to state with clarity any property differentiations. As a matter of fact all gradations of behavior ranging from pure covalency on the one hand to almost electrovalency on the other may be found. In general the evidence seems to indicate that coordinate covalent binding tends to increase the melting and boiling points of compounds containing only covalent bonds.

Exceptions to the Octet Rule. The Lewis-Kossel theory through its postulate of a tendency toward a rare gas configuration in chemical combination can and has accounted for the formation of many compounds by electron transfer and sharing. In doing this the theory relied on the concept that for stability there can be present no more than eight electrons about an atom. However, many compounds are known for which this maximum electron octet rule is apparently exceeded, namely, PCl$_5$, SF$_6$, OsF$_8$, etc. In terms of an electron pair per bond, the phosphorus atom in PCl$_5$ would have to be surrounded by 10 electrons, the sulfur in SF$_6$ by 12, and the osmium in OsF$_8$ by 16. Again, in a molecule such as B$_2$H$_6$ there are present insufficient electrons to bind by sharing both the two borons and the six hydrogens. Since each boron possesses three electrons, the total of six for the two would be just sufficient to bind the hydrogens without leaving any for the bond between the borons. Still, although not very stable, B$_2$H$_6$ exists.

To account for the existence, and in some instances high stability, of compounds apparently violating the octet rule, the suggestion has been made that more than eight electrons may be coordinated about an atom. However, Sugden pointed out that the octet rule can be preserved provided it is granted that in some cases a bond can be formed by the

sharing of a single electron rather than a pair. Thus for phosphorus pentachloride the structure may be written as a combination of three electron pairs and two singlet electron bonds, namely,

$$
\begin{array}{c}
\text{Cl} \\
\;\; \vdots \;\;\; \text{Cl} \\
\text{Cl} : \ddot{\text{P}} : \\
\;\; \vdots \;\;\; \text{Cl} \\
\text{Cl}
\end{array}
$$
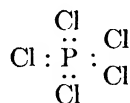
for SF_6 as a combination of two pairs and four singlets, while for OsF_8 of eight singlet bonds. Likewise the formation of B_2H_6 may be accounted for by the structure

$$
\begin{array}{c}
\text{H} \quad \text{H} \\
\;\; \vdots \;\; \vdots \;\; \\
\text{H} : \ddot{\text{B}} : \ddot{\text{B}} : \text{H} \\
\;\; \cdot \;\; \cdot \;\; \\
\text{H} \quad \text{H}
\end{array}
$$
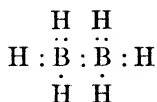
Although there seems to be little doubt that single electron bonds are possible under certain conditions, as in the hydrogen-molecule ion, $[H \cdot H]^+$, found in discharge tubes containing hydrogen, it is questionable whether they exist in the form postulated in all such molecules. Where they do exist the stability results apparently from a quantum-mechanical phenomenon called *resonance*, in which the singlet electron, instead of being fixed, oscillates between bonds involving the singlets. For example, in B_2H_6 each singlet electron would oscillate between one hydrogen and another, with the result that each singlet bond would possess an electron doublet for part of the time of an oscillation. As quantum or wave mechanics has shown that whenever resonance appears it leads to increased stability, the formation of stable bonds with supposedly singlet electrons may well be the consequence of resonance.

Wave-Mechanical Theories of Valence. Two theories of valence based on wave mechanics have been advanced which attempt to express the concept of valence and chemical combination in quantitative terms. Although both of these theories, the *Heitler-London* theory and the *Hund-Mulliken* theory, employ the same principles, they differ in mode of treatment. The Heitler-London theory ascribes particular electrons to the atoms from which they come, and considers the manner of interaction of these with other electrons during chemical combination. On the other hand, the Hund-Mulliken theory does not treat the electrons as belonging to a particular atom, but considers rather the molecule as a whole and the manner in which all the electrons present interact with each other and with the positive atomic nuclei. Despite this difference of approach, both theories lead to conclusions which in all essential details are the same. In this connection it may be emphasized that, because of mathematical difficulties, exact solutions of the wave equa-

tions involved have been obtained only in such simple systems as the hydrogen molecule and the hydrogen-molecule ion, H_2^+. In more complicated cases various approximate methods of solution must be resorted to. Nevertheless, even these approximate treatments yield considerable information on valence and bonding, and they are, therefore, of great interest and importance to the chemist.

According to the Heitler-London theory the valence of an element results from the presence in the external shell of the atom of electrons with *unpaired spins*. By an electron with unpaired spin is meant an electron with spin quantum number $+ 1/2$ or $- 1/2$ for which no complementary electron with *opposite* spin is present in the atom. As long as electrons with opposite spins are paired off, they are not available for chemical combination. When they are not paired off, they may satisfy their tendencies for spin coupling by combining with the electrons of opposite spin in other atoms to form a valence bond and thereby a molecule. *Valence and chemical combination are the result, therefore, of the tendency of unpaired electrons with opposite spins to pair off and balance thereby the electronic spins.* It will be observed that the electron pair covalent bond concept of Lewis finds thus its explanation in this theory as a coupling of two electrons with opposite spins.

The Heitler-London theory emphasizes the electron pair type of bonding, i.e., covalency, and so it is frequently referred to as the *homopolar method*. On the other hand, the Hund-Mulliken or *molecular orbital* treatment results not only in homopolar wave function terms in the wave equation of the molecule, but also in certain extra terms corresponding to possible ionic binding. When the values of the homopolar terms are considerably greater than those of the ionic, the binding is essentially of the covalent type. On the other hand, when the ionic terms are much more significant than the homopolar, we encounter electrovalence. Between these two extremes the binding is of mixed type and leads to the appearance of molecules with *permanent dipole moments* to which reference will be made in the next chapter.

Another difference between the two theories lies in the fact that, whereas the homopolar theory emphasizes the pairing of electrons of opposite spins, the molecular-orbital method divides all the electrons in the molecule into *bonding, antibonding,* and *nonbonding* electrons. Any electron whose presence contributes to molecular stability is a bonding electron, while one whose presence leads to instability is an antibonding electron. A nonbonding electron, in turn, is one which contributes neither to stability nor instability. Of necessity any stable molecule will be one containing an excess of bonding over antibonding electrons. When the reverse is true, the molecule cannot be stable, and chemical combination of the atoms involved is not possible.

Resonance. A fundamental contribution of wave mechanics to the theory of molecular structure is the concept of *resonance* or *exchange energy*. This idea of resonance has no counterpart in any classical theory of atomic combination. In order to explain it let us consider as a specific instance the formation of a hydrogen molecule from two hydrogen atoms. Each of these hydrogen atoms contains one electron. When the molecule is formed it may be thought off-hand to consist of the two atoms bound together in a manner such that each electron remains essentially attached to its own atom. But it is quite possible that the two electrons may exchange orbits, namely, electron (1) from atom (1) may exchange places with electron (2) from atom (2). When this happens the molecule may consist not only of the configuration atom (1) electron (1)—atom (2) electron (2), but also of atom (1) electron (2)—atom (2) electron (1). These two configurations are not static, but involve *continuous oscillation* of the molecule from one of these to the other. Wave mechanics shows that, when such electron exchange is possible and does occur, the energy of the resonating structure is *less* than that of the molecule without exchange; i.e., that *exchange or resonance leads to an increase in stability*. The increase in binding energy thus resulting is called *resonance* or *exchange energy*.

Resonance energy may arise not only from electron exchange as described, but also from the oscillation of electrons from one position in the molecule to another. As a consequence of this oscillation the molecule may take on a number of resonance structures, each of which is essentially equivalent in total energy to another. Thus the carbon dioxide molecule may have the possible equivalent structures

$$O : C : : : O \qquad O : : C : : O \qquad O : : : C : O$$
$$\text{I} \qquad\qquad\qquad \text{II} \qquad\qquad\qquad \text{III}$$
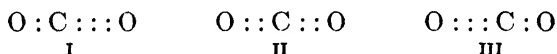
At one time structure II was thought to be predominant, with I and III practically nonexistent. However, at present it is believed that carbon dioxide is a resonance hybrid of all three forms; and, as a consequence, its stability is greater than it would be if only one structure persisted. Similarly it has been shown that the stability of the benzene molecule is due to a large degree to resonance between at least five equivalent structures. In fact, whenever the possibility of resonance exists, greater stability in a molecule or ion can be anticipated.

The Hydrogen Molecule in Wave Mechanics. Since the formation of a hydrogen molecule from two hydrogen atoms has been treated exactly by wave mechanics, it is of some interest to recapitulate briefly the results obtained. A solution of the appropriate wave equations indicates that the total energy of the hydrogen molecule consists of three parts: (1) the energy of the original atoms; (2) coulombic energy result-

ing from electrostatic attractions between the electrons and nuclei; and (3) exchange energy. Further, the treatment leads to two expressions for the total energy as a function of the distance of separation in the two nuclei. In one of these the total energy is always such as to lead to continual *repulsion* of the two atoms at all possible distances of separation, and hence this solution cannot possibly correspond to the formation of a stable molecule. This case is illustrated by curve A in Fig. 1. For this plot the ordinate is the total energy of the molecule taken on the basis that the energy of the atoms is zero at infinite distance of separation, while the abscissa is the distance of separation of the two nuclei, r. In turn, the second solution gives curve B for the total energy as a function of the distance of separation. This curve shows that as the two nuclei are brought closer together an attraction results which increases as r decreases, until at $r = r_0$ the attraction is a maximum. r_0 must correspond, therefore, to the distance of separation of the two nuclei in the stable molecule. If this distance is decreased, the plot shows that the attraction is also diminished, until at very small distances of separation the two nuclei will actually repel each other.

A plot such as Fig. 1 of energy vs. distance of separation of the nuclei is called a *potential energy diagram*. The potential energy curve for the hydrogen molecule indicates that maximum attraction between two hydrogen atoms will occur at a distance r_0 apart, and hence at this stage we shall have the normal hydrogen molecule. To disrupt the hydrogen molecule into its constitutent atoms it would be necessary to
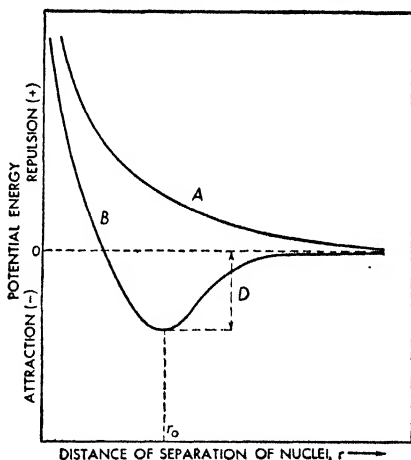
Fig. 1. Potential Energy Diagram of H_2 Molecule

increase the distance of separation r theoretically to infinity. For this purpose an amount of energy would be required corresponding to the difference between the energy of the molecule D and the energy of the two atoms, which is zero by definition. Consequently the energy D must be the *energy of dissociation* of a hydrogen molecule into its constituent atoms. Spectroscopy yields for this energy the value $D = 102,900$ calories per gram mole, corresponding to a value of $r_0 = 0.74 \times 10^{-8}$ cm. The values of these two constants for the hydrogen molecule deduced by James and Coolidge from wave mechanics are exactly the same, and hence theory and experiment are here in exact accord.

The Hydrogen Bond. Before concluding the discussion of bonds and bonding, it is of interest to point out that in certain compounds a hydrogen already bound by two electrons to an atom may further coordinate two more electrons to form another bond. Examples of such hydrogen bonding may be found in

water,

$$\text{H—O—H---O}\begin{matrix}\diagup\text{H}\\\diagdown\text{H}\end{matrix}$$

alcohols,

$$\underset{\overset{|}{\text{H}}}{\text{R—O}}\text{---HOR}$$

carboxylic acids,

$$\text{R—C}\begin{matrix}\overset{\text{O---H—O}}{\diagup\!\!/}\quad\diagdown\\\diagdown\quad\underset{\text{O---H—O}}{/\!\!/}\end{matrix}\text{C—R}$$

nitrophenols,

$$\begin{matrix}\diagup\text{O—H}\\\qquad\overset{|}{\underset{|}{\,}}\\\diagdown\text{N}\!\!=\!\!\text{O}\\\overset{|}{\text{O}}\end{matrix}$$

and in amides,

$$\text{R—C}\begin{matrix}\overset{\overset{\text{H}}{|}}{\text{O---H—N}}\diagdown\\\diagdown\quad\underset{\underset{\text{H}}{|}}{\underset{\text{N—H---O}}{}}/\!\!/\end{matrix}\text{C—R}$$
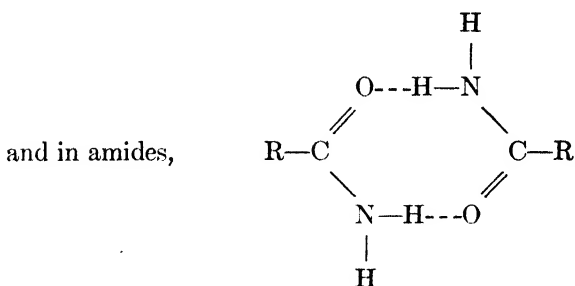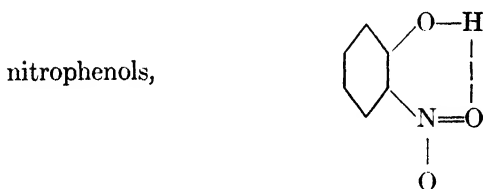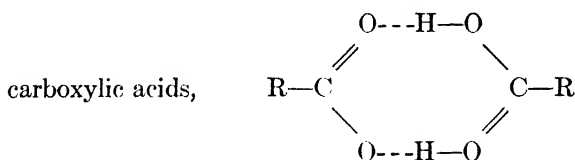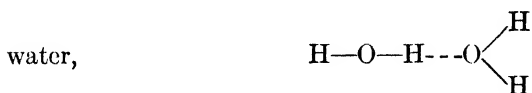
In each of these instances the hydrogen bond arises from the donation of a pair of electrons by a nitrogen or oxygen atom to a hydrogen already bound to a nitrogen or oxygen. In the nitrophenols the hydrogen bonding results in the formation of a second ring in which the properties of both the nitro and hydroxyl groups are modified. Again, with carboxylic acids and amides the hydrogen bonding leads to an association involving two molecules. On the other hand, in water and the alcohols the hydrogen bonding need not stop at two molecules, but may well lead to a chain extending throughout the entire mass. In terms of such

hydrogen bonding it is readily understandable why water, acids, alcohols, and amides are associated liquids, whereas hydrocarbons, carbon tetrachloride, and chloroform are not. The same explanation applies to liquid hydrogen fluoride, where as a result of hydrogen bond formation between the fluorine of one molecule and the hydrogen of another the dimer H_2F_2 is formed.

Hydrogen bonding is not confined to the liquid state. X-ray studies of ice and solid carboxylic acids, and vapor density measurements on the latter, reveal that hydrogen bonding extends also into the solid and vapor phases. The presence of the hydrogen bonds is indicated in the liquid state by boiling points higher than may be anticipated for normal covalent liquids, and by high heats of vaporization. As a general rule the viscosity is also abnormally large. Finally, hydrogen bonding can be detected also by infrared absorption studies. For example, the infrared spectrum of aceto-acetic acid ester shows the absence of bands associated with the hydroxyl group, and this can mean only that this group is tied up by formation of hydrogen bonds. Again, in other molecules the position of the hydroxyl bands is displaced in the infrared, and hence such a displacement may be used to ascertain presence of hydrogen bonding.

The relatively high stability of hydrogen bonds is ascribed to the possibility of existence of resonance in compounds where the bond occurs. Thus in water, the alcohols, and carboxylic acids the hydrogen bonded between the two oxygens is believed to resonate between these two atoms. The same applies to the hydrogen in the nitrophenols. In the amides, on the other hand, the resonance involves alternate attachment of the hydrogen to the oxygen or nitrogen atom, and as a consequence the net structure resulting from the two alternate resonating forms is stable.

Molecular or Band Spectra. Just as study of atomic spectra has proved helpful in elucidating the structure of atoms, so the spectra emitted by gaseous molecules have yielded valuable information on the structure of molecules. Unlike the *line* spectra emitted by atoms, molecules of a gas when excited emit a *band* spectrum which at first sight may appear continuous. However, strong resolution reveals that it is actually composed of many lines very closely spaced. In general the sorting and classification of lines in a molecular band spectrum is a difficult and tedious task. Nevertheless, the spectra of most diatomic and many polyatomic molecules have been unraveled, and from these a deep insight has been obtained into what occurs within the molecule.

Three types of molecular spectra are recognized, namely, (1) *rotation spectra*, (2) *vibration-rotation spectra*, and (3) *electronic band spectra*. To understand the origin of these various spectra an understanding is required of the energy levels in a molecule. For this purpose consider

a diatomic molecule AB. This molecule consists of two nuclei corresponding to the atoms A and B and their accompanying electrons. As in the constituent atoms, the various electrons in the molecule may exist in a number of energy levels such that on excitation an electron may shift from one to another. Once thus excited the electron on return will emit a spectral line of a frequency determined by the energy difference of the two levels and the quantized Bohr relation. Were this the only manner in which the molecule could absorb energy, the electronic spectrum of a molecule would be a line spectrum akin to that of atoms. But a molecule may also absorb energy to cause a *vibration* of the nuclei with respect to each other, and to set the molecule as a whole *rotating*. These vibrational and rotational motions are quantized, and are superimposed upon the kinetic motion of the molecule. From the standpoint of band spectra the latter motion may be disregarded, but the other two motions cannot be. The incidence of quantized vibration and rotation coupled with the electronic levels of the electrons leads to the diagrammatic picture of the energy levels of a molecule shown in Fig. 2. The group of energy levels represented by $n = 1$ corresponds to the lowest electronic energy level of the molecule. Similarly the group given by $n = 2$ is the energy structure of the next higher, excited, electronic level. Each of these electronic levels is subdivided into the

Fig. 2. Molecular Energy Levels (Schematic)

vibrational sublevels designated by the vibrational quantum numbers v on the left. In turn, each of the vibrational sublevels has associated with it a number of *rotational* quantum levels the energy of which is determined by the rotational quantum numbers J shown on the right. In a like manner every higher electronic level of a molecule will consist of vibrational and rotational levels of the form shown in Fig. 2, and will lie in sequence above those for $n = 2$.

From this description it is evident that the energy required for excitation of various emissions will be least for rotation, and will be increasingly greater for vibration and then electron transfer. If the exciting energies are kept suitably small it is possible to cause the excitation of

only rotation, i.e., transitions from one rotational quantum level to another within a given vibrational level, and the result is then the appearance of the *rotational spectrum alone*. Since the energies involved are small so are the frequencies of emitted lines, and hence these spectra are found in the far infrared, i.e., at frequencies of 50 to 100 μ.[1] On the other hand, when the exciting energies are sufficiently large to cause vibration level transitions within a given electronic level, emissions are observed corresponding to the changes in the vibration quantum numbers. Further, since a change in vibrational levels involves also changes in rotational levels, the overall result is a vibrational spectrum wherein each line is accompanied by a rotational fine-structure. We obtain thus the *vibration-rotation spectrum* of a molecule which is generally located in the near infrared, i.e., at frequencies of 1 to 50 μ. Finally, with still higher exciting energies electron transitions may take place which are accompanied by vibrational level changes, and each of these in turn by rotational fine-structure. The spectrum observed is then a complex *electronic band* consisting of lines due to electronic, vibrational, and rotational transitions. As more than a single electronic excitation may occur, the complete spectrum may consist of a *band system* composed of individual electronic bands, each accompanied by its own vibrational and rotational lines. Electronic bands are found in the visible and ultraviolet spectral ranges.

Rotational Spectra. The rotational spectrum of a *diatomic* molecule can be accounted for by considering the molecule to a first approximation as a rigid rotator, i.e., as a rigid dumbbell joined along its line of centers by a bond equal in length to the distance r_0 between the two nuclei. With the energy of rotation quantized, wave mechanics shows that the energy E_J of such a rigid rotator in any given rotational quantum level J is given by

$$E_J = \left(\frac{h^2}{8\pi^2 I}\right) J(J+1) \tag{1}$$

In this equation h is again Planck's constant, while I is the *moment of inertia* of the molecule given in terms of the two atomic masses, m_1 and m_2, and the internuclear distance r_0 by the relation

$$I = \left(\frac{m_1 m_2}{m_1 + m_2}\right) r_0^2 \tag{2}$$

J may have integral values of 0, 1, 2, 3, etc. For a transition from a rotational level of quantum number J to one of lower quantum number J', the energy difference would correspond to

$$\Delta E_r = E_J - E_{J'} = \left(\frac{h^2}{8\pi^2 I}\right)[J(J+1) - J'(J'+1)] \tag{3}$$

[1] $1\,\mu = 10^{-4}$ cm $= 10,000$Å.

As a rule rotational transitions are restricted to changes in J corresponding to $\Delta J = J - J' = 1$. Inserting this condition into equation (3) and remembering that the frequencies of the resulting spectral lines are given by $\Delta E_r = h\nu$, we obtain

$$\Delta E_r = h\nu = \left(\frac{2\,h^2}{8\,\pi^2 I}\right) J \tag{4}$$

and hence

$$\nu = \left(\frac{h}{4\,\pi^2 I}\right) J \tag{5}$$

In this equation the values of J start with $J = 1$. Since the quantity in parentheses is a constant for a given molecule, and as J has integral values, equation (5) shows that the rotational spectrum should consist of a number of equally spaced lines the frequencies of which are determined by the value of J. Hence, by determining the frequencies of lines corresponding to various values of J in the pure rotation spectrum of a molecule, the moment of inertia of the molecule may be obtained through equation (5), and from it the internuclear distance of the atoms r_0 by means of equation (2).

Although pure rotation spectra offer apparently the simplest means of deducing moments of inertia of molecules, technical difficulties in far infrared studies impair high accuracy. For this reason this quantity is deduced more frequently from the other types of spectra, which, despite their higher complexity, can be measured more conveniently and precisely.

Vibration-Rotation Spectra. To a first approximation the vibratory motion of the nuclei of a diatomic molecule can be represented as the vibration of a *simple harmonic oscillator;* i.e., an oscillator in which the restoring force is proportional to the displacement in accordance with Hooke's law. For such an oscillator with quantized energy, wave mechanics shows that the vibrational energy E_v is related to the fundamental vibrational frequency ν_0 by the relation

$$E_v = \left(v + \frac{1}{2}\right) h\nu_0 \tag{6}$$

Here v is the *vibrational quantum number*, which may take on the values $v = 0, 1, 2$, etc. Equation (6) reveals that such an oscillator, even in its lowest vibrational level, $v = 0$, retains the energy $E_0 = 1/2\,h\nu_0$. This residual vibrational energy is called the *zero-point energy* of the oscillator, and it cannot be removed from the molecule even by cooling it to the absolute zero.

For a vibrational transition from any level of quantum number v to another of quantum number v', the emitted energy is given by

$$\Delta E_v = E_v - E_{v'} = \left(v + \frac{1}{2}\right) h\nu_0 - \left(v' + \frac{1}{2}\right) h\nu_0 = (v - v')h\nu_0 \tag{7}$$

As was pointed out, upon these vibrational transitions are superposed accompanying rotational changes which may *add* or *subtract* from equation (7) the energy change given by equation (4). The energy change resulting is, therefore, the *sum* or the *difference* of equations (7) and (4), or

$$\Delta E = \Delta E_v \pm \Delta E_r$$

$$= (v - v')h\nu_0 \pm \left(\frac{h^2}{4\,\pi^2 I}\right) J \tag{8}$$

The emitted spectral lines due to the energy change given by equation (8) follow from $\Delta E = h\nu$. Hence,

$$\nu = \frac{(v - v')h\nu_0 \pm \left(\dfrac{h^2}{4\,\pi^2 I}\right) J}{h}$$

$$= (v - v')\nu_0 \pm \left(\frac{h}{4\,\pi^2 I}\right) J \tag{9}$$

From equation (9) it can be deduced that the appearance of a vibration-rotation spectrum will be as follows: Instead of obtaining a single line corresponding to the vibrational transition from a given value of v to v', the line will be missing, and instead a pair of lines on either side of the expected position will be found for each value of J, one corresponding to the plus sign in equation (9), the second to the minus sign. Since a pair of such lines will be obtained for each of the possible values of J, i.e., $J = 1, 2$, etc., the spectrum for the particular difference in v will have a blank space flanked on either side by the rotational fine-structure
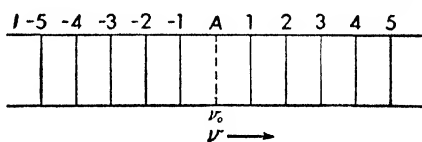
Fig. 3. Diagram of Vibration-Rotation Spectrum (Schematic)

lines. Such a spectrum is shown schematically in Fig. 3. There the spacings between the lines are given as constant in line with equation (9), but actually the spacings are variable because of complications to be mentioned later. As one unit such as Fig. 3 is obtained for each possible change in vibrational quantum number from v to v', the complete vibration-rotation spectrum will be composed of a number of such units.

The bands corresponding to $v' = 0$ and $v = 1, 2$, etc., are called respectively the *fundamental* band, the *first harmonic* band, the *second harmonic* band, etc. By ascertaining in the fundamental band the frequency corresponding to the missing vibration line, i.e., position A in Fig. 3, ν_0 is obtained directly. Similarly the frequencies at the positions corresponding to A in the first and second harmonics give $2\,\nu_0$ and $3\,\nu_0$. These

may be used, then, to check the value of ν_0, the fundamental vibration frequency of the molecule. With ν_0 thus known, measurement of the line frequencies for various values of J permits calculation of the moment of inertia through equation (9), and from it, in turn, the internuclear distance r_0. In this manner it is possible to ascertain from vibration-rotation spectra the fundamental frequency of the molecule, the moment of inertia, and the distance of separation of the nuclei.

Electronic Spectra. Electronic spectra arise from excitation of electrons in a molecule from one energy level to a higher one. As a result of this excitation the electron on return will emit radiation corresponding to the energy difference between the two electronic levels involved. Since during the excitation a change takes place not only in electronic but also in vibrational and rotational levels, the electronic emission is always accompanied by vibrational and rotational transitions. Thus, if we let E'_e, E'_v, and E'_r be, respectively, the electronic, vibrational, and rotational energies of the molecule after the transition, and E_e, E_v, and E_r the corresponding quantities before the change, then the total energy of the molecule in the initial state is

$$E = E_e + E_v + E_r \tag{10}$$

and, in the final state,

$$E' = E'_e + E'_v + E'_r \tag{11}$$

The energy change involved in the electronic transition is, consequently,

$$\begin{aligned} \Delta E &= (E_e - E'_e) + (E_v - E'_v) + (E_r - E'_r) \\ &= \Delta E_e + \Delta E_v + \Delta E_r \end{aligned} \tag{12}$$

and the resulting frequency of radiation follows as

$$\nu = \frac{\Delta E}{h} = \frac{\Delta E_e + \Delta E_v + \Delta E_r}{h} \tag{13}$$

As ΔE_v and ΔE_r may have different values depending on the rotational and vibrational quantum numbers involved, the number of lines possible for a given change in E_e is large and leads to a complex electronic band. Further, with different values of ΔE_e possible as well, the complete band spectrum will consist of a band system composed of a series of individual bands.

Despite their complexity electronic band spectra are frequently studied in preference to the other types. This is because these spectra occur in the visible or ultraviolet ranges and hence may be photographed and inspected easily. In general the information deduced from electronic bands is the same as that obtained from vibration-rotation spectra.

Besides, it is possible to ascertain from electronic band spectra information about the excited states of the molecule, as well as the energy of dissociation of the molecule into its component atoms.

When the electronic band spectrum of a molecule is inspected, it is found that at one end of the spectrum the distance of separation of the lines becomes smaller and smaller, until at a given position the line spectrum terminates, and a space of continuous emission starts. The position of the limit where the lines end and the continuum begins gives the frequency of radiation necessary for *dissociation* of the molecule. When the dissociation is into normal atoms the spectroscopic dissociation energy D_s can be calculated directly from the observed frequency. When the dissociation results in one or more excited atoms, however, a correction has to be applied for the energy of excitation. From the value of D_s thus obtained the thermochemical dissociation energy of the molecule follows then as

$$D = D_s + \frac{1}{2} h\nu_0 \tag{14}$$

where the second term on the right gives the zero-point energy of the molecule. For any diatomic molecule, such as hydrogen, D given by equation (14) is identical with the energy D indicated in Fig. 1.

The above discussion of the molecular spectra of diatomic molecules is a simplification of the actual situation. The molecule as a whole is not quite a rigid rotator nor a simple harmonic oscillator. Because of these defects in ideal behavior corrections have to be introduced which tend to complicate equations (5), (7), and (9). For the purpose at hand, however, the treatment given should suffice.

In Table 2 are summarized the molecular constants of some diatomic molecules deduced from band spectra in the manner described. In this

<div align="center">

TABLE 2

MOLECULAR CONSTANTS OF DIATOMIC MOLECULES

</div>

Molecule	$m^* \times 10^{24}$ (g)	$I \times 10^{40}$ (g-cm²)	$r_0 \times 10^8$ (cm)	$\nu_0 \times 10^{-12}$ (sec⁻¹)	D Kcal/mole	D Thermal Kcal/mole
H_2	0.832	0.46	0.75	132.4	102.9	103.2
N_2	11.54	13.8	1.09	70.76	170.	
O_2	13.19	19.15	1.20	47.42	116.7	
Cl_2	28.86	113.5	1.98	16.03	56.8	57.6
Br_2	65.89	445.	2.28	9.71	45.2	46.2
I_2	104.65	740.	2.66	6.42	35.4	35.1
HCl	1.62	2.65	1.28	87.70	102.	107.0
CO	11.31	15.0	1.13	65.07	223.	
NO	12.31	16.4	1.15	57.18	123.	

table column 2 gives the reduced mass of the molecule, i.e., $m* = (m_1 m_2)/(m_1 + m_2)$, columns 3, 4, and 5 the moments of inertia, the interatomic distances, and the fundamental frequencies of vibration, respectively, while column 6 gives the spectroscopically deduced dissociation energies. For comparison there are included in column 7 some values of the dissociation energy obtained by thermochemical methods. The agreement between the two sets of D values is satisfactory.[1]

Potential Energy Diagrams of Diatomic Molecules. It was indicated before that quantum mechanics leads to the potential energy-distance of separation diagram for the stable hydrogen molecule depicted by curve B in Fig. 1. A similar type of diagram can be constructed for every diatomic molecule with the aid of the molecular constants deduced from spectroscopy. For the *lowest* or *ground electronic state* of a molecule the potential energy diagram will have the general form shown in Fig. 4. To any higher or excited state an analogous potential energy curve will correspond, except that because of the higher energy content it will lie above the one for the ground state.

To understand Fig. 4 it must be realized that the total energy E of any molecule consists of the electronic energy E_e, the vibrational energy E_v, the rotational energy E_r, and the energy of thermal motion E_t; i.e.,

$$E = E_e + E_v + E_r + E_t \qquad (15)$$

At the absolute zero of temperature E_t and E_r are both zero, and E_v reduces to the zero-point energy of the molecule, namely, $E_0 = 1/2\, h\nu_0$. At the absolute zero, therefore, the energy is given by

$$E_{T=0} = E_e + \frac{1}{2}\, h\nu_0 \qquad (16)$$

For such a stable molecule in its ground state point S in Fig. 4, corresponding to the normal distance of separation r_0, gives the electronic energy content E_e. When the zero-point energy is added to this quantity, the energy level is raised to the line A–A, and we arrive thus at the energy of the molecule at the absolute zero. As the vibrational frequencies of the molecule are excited, the total energy levels are raised to the lines B–B, C–C, etc., in accordance with the increase in the vibrational quantum number v. The spacings of the lines within the potential energy diagram depend on the differences in energy between vibrational quantum levels. For an ideal harmonic oscillator the

---

[1] In terms of ordinary thermodynamic quantities, D is the internal energy change on dissociation of a diatomic molecule into atoms at the *absolute* zero of temperature, i.e., $D = \Delta E_0$. To obtain the value of ΔH from D at any temperature T the thermal energy entering between $T = 0$ and $T = T$ must be added to ΔE_0. In other words, $\Delta H_T = \Delta E_0 + \int_0^T \Delta C_p dT$. The thermal values of D given in column 7 are corrected to $T = 0$.

spacings between the lines would be constant, but actually they get smaller as v increases. When eventually line X–X is reached, the molecule dissociates and we obtain thus the two neutral atoms. The difference in energy between the level X–X and $E = 0$, namely E_a, is of necessity the energy content of the two resulting atoms. If from this
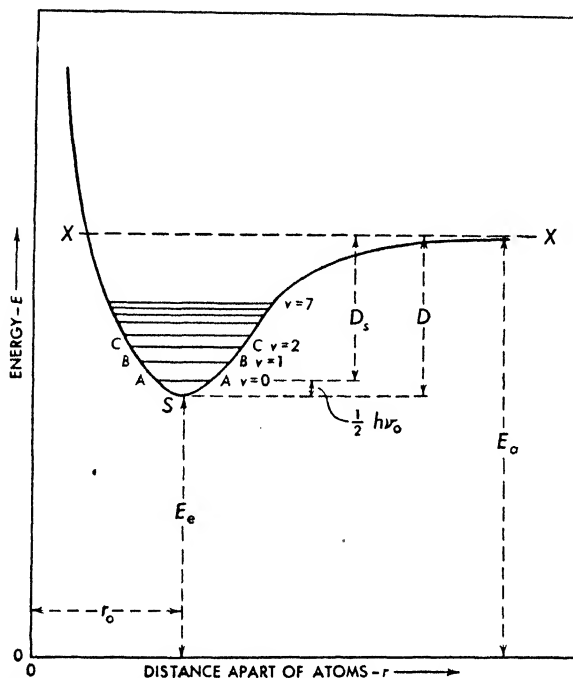
Fig. 4. Potential Energy Diagram of Diatomic Molecule

energy we subtract the electronic energy E_e of the original molecule, the result is the energy of dissociation D of the molecule. However, if instead of subtracting E_e we deduct $E_e + 1/2\, h\nu_0$, we obtain the spectroscopic energy of dissociation D_s.

A diagram such as Fig. 4 indicates, therefore, the energy of a molecule in its ground electronic state and in various vibrational states as a function of the distance of separation. To obtain the total energy of the molecule at any given temperature it would be necessary to add to $(E_e + E_v)$ the rotational and thermal energy contributions.

In dealing with the thermal properties of molecules, such as specific heats and other thermodynamic properties, the consideration is always with energy differences, and hence in absence of electronic excitation the electronic energy cancels out. Under such conditions E_e may be disregarded, and the energy reference point may be transposed to point

S; i.e., the energy at S may be taken as zero. On this basis the potential energy E_P of any diatomic molecule can be derived from an empirical equation due to Morse, namely,

$$E_P = D[1 - e^{-a(r-r_0)}]^2 \qquad (17)$$

Here again D is the energy of dissociation and r_0 the equilibrium separation distance of the molecule, r is the distance of separation at any other point, and a is a constant characteristic of a particular molecule. The constant a may be evaluated from the relation

$$a = \pi\nu_0 \sqrt{\left(\frac{2\,m_1m_2}{m_1 + m_2}\right)\frac{1}{D}} \qquad (18)$$

where ν_0 is the fundamental vibration frequency of the molecule, while m_1 and m_2 are the masses of the two atoms involved. Therefore, knowing D and ν_0 for a molecule, a may be found through equation (18), and therefrom and r_0 the potential energy through equation (17). In using equation (18) D must be expressed in *ergs*, but in equation (17) it may have any desired units, as kilogram-calories. Finally, to establish the vibrational levels, i.e., the lines A–A, B–B, etc., in Fig. 4, recourse is had to equation (6). By substituting into this relation $v = 0$ the zero-point energy follows, $v = 1$ the level B–B, $v = 2$ the level C–C, etc. The lines thus obtained will be idealized and equidistant. For their true positions information on the anharmonicity of the oscillator deduced from actual data is required.

Polyatomic Molecules. In describing band spectra and their origin attention was confined to diatomic molecules. With molecules containing more than two atoms the situation is more involved. Any diatomic molecule must of necessity be linear, and hence it can have only one moment of inertia. This is true also of linear polyatomic molecules. However, when the polyatomic molecules are nonlinear, i.e., when the valence bond directions are oriented at angles other than 180° to each other, then *three* moments of inertia are present. These three moments of inertia correspond to the respective rotations of the molecule as a whole about a set of x, y, and z axes passed through the center of gravity of the molecule as an origin. Again, whereas a diatomic molecule can have only one fundamental frequency of vibration, the number of such fundamental frequencies increases with the number of atoms. It can be shown that for a linear polyatomic molecule the number of possible modes of vibration the molecule *may* have is $(3\,n - 5)$, where n is the number of atoms in the molecule, while for a nonlinear molecule it is $(3\,n - 6)$. Thus, in passing from a diatomic to a triatomic linear molecule the possible vibration frequencies increase from one to four, and for a nonlinear molecule from one to three. Not all the possible vibrations

are necessarily active, nor are all of the frequencies necessarily different. Nevertheless, the added modes of motion complicate the situation, and consequently the resulting spectra are much more complex and difficult to interpret.

Raman Spectra. An alternate and simpler method than band spectroscopy of obtaining vibrational and rotational frequencies of molecules is through observations of the *Raman effect*. In 1928 Raman found that when light of a definite frequency ν was passed through a gas, liquid, or solid, and the scattered radiation was observed at right angles to the direction of the incident beam, the scattered radiation contained lines not only of the original light, but also some of lower, and occasionally higher frequency. The lines shifted toward the lower frequencies are called *Stokes lines*, those toward the higher frequencies *anti-Stokes lines*. Raman further found that the difference $\Delta\nu$ between the incident and any given scattered line was *constant and characteristic of the substance irradiated*, and *completely independent of the frequency ν of the incident radiation*. Such differences $\Delta\nu$ between incident and scattered lines are called *characteristic Raman frequencies*. Evidently the characteristic Raman frequencies must be due to internal energy changes occurring within the molecule as a result of interaction with the incident light, and must find their explanation in terms of these.

It is now known that the Raman effect arises from the absorption by a molecule from the incident radiation of sufficient quanta of energy to cause transitions from lower to higher vibrational levels, or from lower to higher rotational levels. In vibrational excitation a *single* quantum of energy is removed in magnitude equal to the energy necessary for passage from one vibrational quantum level to the next. In rotations, however, *two* quanta are extracted to cause a change in rotational level from a value J to $(J + 2)$. As a result of such removal of energy from the original beam the energy of the latter must decrease, and so must the frequency. We obtain thus the Stokes lines. On the other hand, the molecule, instead of absorbing energy, may emit some by rotational or vibrational transitions from higher levels to lower. When this happens the emitted energy is added to that of the incident light, and hence the frequency of the scattered light is greater. The results then are anti-Stokes lines.

From the above explanation it is to be anticipated that the Raman frequencies observed for various rotational and vibrational transitions should be identical with the same transitions observed in band spectra. This is the case, as may be judged from Table 3, where Raman and band vibrational frequencies of several molecules are compared. Since this is true, vibrational and rotational frequencies may be ascertained more conveniently from the Raman effect, which may be studied in the

TABLE 3

COMPARISON OF RAMAN AND BAND SPECTRAL
VIBRATIONAL FREQUENCIES

Molecule	$\Delta\nu$ Raman (cm^{-1})	ν_0 from Band Spectra (cm^{-1})
O_2	1555	1554
N_2	2331	2331
H_2	4160	4159
CO	2145	2144
HCl	2886	2886
HBr	2558	2559

visible spectral range rather than in the more difficult infrared, where the rotational and vibrational bands are found. Utilizing this advantage, investigators have identified and measured the frequencies corresponding to various types of bonds in complex organic and inorganic molecules. All in all, Raman spectroscopy has proved a great aid in the study of molecular structure and has supplied a check on the interpretation of the band spectra of molecules. It should be pointed out, however, that as a rule more lines may be observed in the infrared than in the Raman spectrum. On the other hand, certain molecules such as O_2, N_2, and H_2 exhibit a Raman spectrum, but they do not yield infrared spectra.

Ortho- and Para-Hydrogen. In discussing the structure of atoms it was pointed out that electrons in atoms revolve about their own axes and they possess, therefore, quantized spin. A detailed analysis of the spectra of atoms reveals also that not only may the electron possess spin but the nucleus as well. The units of quantized spin which a nucleus may possess may be zero for no spin, or a whole number multiple of one-half; namely, 1/2, 1, 3/2, etc. For any given atomic nucleus the spin quantum number is fixed. Thus the nuclei of atomic oxygen, carbon, or helium have no spin, and hence for these the nuclear spin quantum numbers are zero. On the other hand, the spin quantum numbers for the nuclei of hydrogen and fluorine are 1/2, for deuterium and nitrogen 1, for sodium 3/2, and for chlorine 5/2.

Now, when *two identical atoms with spin* are combined into a molecule, as two hydrogen atoms into H_2, two possible combinations arise: (*a*) one in which the two atoms involved have nuclei spinning in the same direction, i.e., *parallel* spins, or (*b*) another in which the two nuclei have opposite or *antiparallel* spins. For the hydrogen nuclei with parallel spins we have what is called *ortho-hydrogen*, and this molecule reveals itself in its rotational band spectrum in fairly intense lines corresponding to rotational levels of *odd rotational quantum numbers J* in the ground state of the molecule. On the other hand, when the spins are antiparallel, the molecule obtained is called *para-hydrogen*, and its rotational band spectrum shows less intense lines corresponding to *even*

rotational quantum levels in the ground state. Band spectra of ordinary hydrogen gas do not fall into either of these categories but show alternate intense and less intense rotational lines. This must mean that ordinary hydrogen is not a single specie, but is in fact a mixture of ortho- and para-hydrogen. From a comparison of the intensities of the rotational lines in ordinary hydrogen it has been established that under ordinary conditions the ratio of ortho- to para-hydrogen is fixed, and is equal to 3 parts ortho to 1 part para in equilibrium with each other.

When a molecule is cooled toward absolute zero, the rotational energy should decrease to zero. Hence it is to be anticipated that near the absolute zero of temperature all the molecules should wind up in the lowest or $J = 0$ rotational level. Since zero may be looked upon here as an even number, this means that near the absolute zero all of the molecules in hydrogen should be converted into para-hydrogen. But, owing to the difficulty of reversing nuclear spins, considerable time is required for the transformation of ortho- to para-hydrogen, and consequently the molecules on cooling may not all enter the lowest rotational level. It has been found, however, that the establishment of the requisite equilibrium at low temperature can be accelerated by the use of charcoal as a catalyst at the temperature of liquid hydrogen. By this means almost pure para-hydrogen has been prepared and studied. It has thus been found to have many properties different from those of ordinary hydrogen. For example, whereas the vapor pressure of ordinary hydrogen at 13.95° K is 53.9 mm, that of para-hydrogen is 57.0 mm Hg. The boiling points, triple points, specific heats, and electrical conductivities are also different. On the other hand, the heats of fusion and vaporization, the gaseous viscosity, and the lattice type and parameters in the solid state are apparently the same. From the standpoint of the properties which are different, as the specific heat, the values observed on ordinary hydrogen are statistical averages of the properties of ortho- and para-hydrogen composing the mixture.

Ortho and para states are observed in molecules only when the two atoms composing the molecule are identical and when they possess nuclear spin. When the nuclei of the atoms do not possess spin, or when the two atoms are different, no ortho-para states can exist. Thus D_2 exists in ortho and para states, but O_2 and HCl do not. In O_2 the nuclei do not have spin. Again, although nuclear spin is present in both H and Cl, the two atoms are different in hydrogen chloride and hence no ortho-para states are possible.

In conclusion it is of interest to point out that existence of ortho and para states increases the number of species of molecular hydrogen to at least five, namely: ortho-H_2, para-H_2, ortho-D_2, para-D_2, and HD. With tritium, $_1^3H$, a possibility, this number may be even greater.

Other Applications of Band Spectroscopic Data. With the aid of statistical mechanics it is possible to derive expressions for the thermodynamic properties of gases, such as the heat content, energy, free energy, heat capacity, and entropy, in terms of the electronic, vibrational, and rotational energies of the molecules. Consequently, as soon as the latter energies are known from band spectra, the thermodynamic properties of the molecules may be evaluated with great accuracy. Although a discussion of the methods involved is beyond our scope, suffice it to say that quite a number of such calculations have been carried out with great success.[1] In limited simple cases band spectroscopic data have also proved valuable in calculating the energies of activation of reactions.[2]

REFERENCES FOR FURTHER READING

1. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.
2. S. Glasstone, "Recent Advances in Physical Chemistry," third edition, The Blakiston Company, Philadelphia, 1936.
3. S. Glasstone, "Recent Advances in General Chemistry," The Blakiston Company, Philadelphia, 1936.
4. G. Herzberg, "Molecular Spectra and Molecular Structure of Diatomic Molecules," Prentice-Hall, Inc., New York, 1939.
5. G. N. Lewis, "Valence and the Structure of the Atom," Reinhold Publishing Corporation, New York, 1924.
6. E. A. Moelwyn-Hughes, "Physical Chemistry," Cambridge University Press, Cambridge, Eng., 1940.
7. O. K. Rice, "Electronic Structure and Chemical Binding," McGraw-Hill Book Company, Inc., New York, 1940.
8. N. V. Sidgwick, "Some Physical Properties of the Covalent Link in Chemistry," Cornell University Press, Ithaca, N. Y., 1933.

PROBLEMS

1. Show the external electronic configuration of the following substances: (a) H_2O_2, (b) CH_3COOH, (c) H_2SO_4, (d) butadiene (C_4H_6), and (e) $HClO_4$.
2. Show the electronic configurations of the following complex ions:

$$(a)\ [Cu(NH_3)_4]^{++}, \quad (b)\ [(CH_3)_4N]^+, \quad and \quad (c)\ [Ag(CN)_2]^-$$

   Which of the bonds are covalent and which are coordinate covalent?
3. If N can have no more than eight electrons in its valence shell, what is the most reasonable electronic formula for R_3NO, where R is an alkyl group?
4. In the CO molecule the binding energy appears to be greater than may be expected for a C–O double bond. Explain what resonating structures of CO might account for this fact.

---

[1] See Taylor and Glasstone, "Treatise on Physical Chemistry," third edition, D. Van Nostrand Company, Inc., New York, 1942, Vol. I, Chapter IV.
[2] See Glasstone, Laidler, and Eyring, "Theory of Rate Processes," McGraw-Hill Book Company, Inc., New York, 1941.

5. Experimental evidence indicates that HCN and NH_3 are associated through hydrogen in the liquid state. Write electronic formulas to show how this association may take place.

6. From the internuclear distance of 1.15 Å for NO, calculate the moment of inertia for the molecule. What spectral frequency in wave numbers will be emitted when NO falls from the first excited rotational state to the ground rotational state? Assume that no changes take place in vibrational or electronic energy.   *Ans.* $I = 1.64 \times 10^{-40}$ g-cm^2

7. Using the data of problem 6, calculate the rotational frequencies for $J = 2, 3$, and 4.

8. From Table 2 calculate the wave length in angstroms of the fundamental, and the first and second harmonic vibration frequencies of the I_2 molecule. What will be the wave length of the fundamental for $J = \pm 1$?
     Ans. For fundamental $\lambda = 46,730$ Å

9. Repeat problem 10 for $J = \pm 2, \pm 3, \pm 4$. What are the corresponding frequencies in wave numbers?

10. What is the energy difference in calories per mole between I_2 in its lowest energy level and I_2 in its first vibrational and rotational state?   *Ans.* 627 cal/mole

11. What is the energy difference in calories per mole between I_2 in its lowest energy level and I_2 in its first excited rotational state, the vibrational level being the same?   *Ans.* 14.3 cal/mole

12. How many principal moments of inertia and fundamental modes of vibration should methyl fluoride have?

13. The heat of dissociation of Cl_2 gas is 56,800 cal/mole. Calculate the work required to separate the two Cl atoms a distance of 0.1 Å from their equilibrium position in the molecule. Repeat the calculation for a distance of 0.03 Å. Use any necessary data given in Table 2.

14. Calculate the zero point energy of HCl in cal/mole.

15. The band spectrum of Br_2 terminates at a wave length of 5107 Å. The energy of the excited atoms formed in the dissociation is 10,400 cal in excess of that possessed by the normal atoms. What is the energy of dissociation of Br_2 in cal/mole?   *Ans.* 45,200 cal/mole

16. Using the molecular constants for I_2 given in Table 2, calculate the potential energy as a function of the internuclear distance of separation, and construct a Morse potential energy curve for I_2.

CHAPTER XXIII

Physical Properties and the Structure of Molecules

Introduction. The purpose of this chapter is to describe certain physical properties of pure substances and solutions, and to show how these are related to the composition and structure of the molecular species involved.

In discussing solutions it was pointed out that the vapor pressure lowering, boiling point elevation, freezing point lowering, and the osmotic pressures of dilute solutions of nonelectrolytes depend solely on the *number* of molecules of solute present and not on their nature. Such properties we have called *colligative*. Besides colligative properties a system may possess also *additive* and *constitutive* properties. Any property of a system is said to be additive when the value of the property for the system is the sum of the values of the corresponding properties of the constituents. Thus the mass of any compound or solution is the sum of the masses of the constituent atoms. On the other hand, a constitutive property is one whose magnitude depends only on the arrangement of the atoms in the assembly. In general we shall find that the properties to be described in this chapter are both additive and constitutive, i.e., they depend both on the number of atoms present in the molecule or solution and their arrangement. From a study of the constitutive effects it should thus be possible to obtain some information about the structure of molecules, and this is, in fact, the case.

The Molar Volume of Liquids. Kopp (1842, 1855) was the first to point out that the molar volume of many liquids *at their normal boiling points* was essentially an additive property of the volumes of the constituent atoms. To explain such observations Kopp postulated that each atom contributes a definite volume to the volume of the molecule, and that this contribution is characteristic only of the given atom and not its arrangement. The preceding statement is known as *Kopp's law of volumes.*

Study of the molar volumes of various liquids has shown that the molar volume is not only an additive property but also a constitutive one. This is revealed by the fact that the contribution to the total volume made by an oxygen atom is either 7.8 or 12.2 cc depending on

740

whether the oxygen is bound by single bonds to two other atoms, as in hydroxyl or ether groups, or by a double bond to a single atom, as in the carbonyl group. However, Kopp's rule does not always yield satisfactory results even when allowance is made for constitutive factors. For this reason this rule cannot be relied upon to give correct molar volumes of liquids.

The Parachor. A much more satisfactory means of comparing molar volumes is embodied in the concept of the *parachor* proposed by Sugden in 1924. In Chapter III it was indicated that one of the more reliable relations for the variation of the surface tension of a liquid with temperature is the McLeod equation, namely,

$$\gamma = C(d_l - d_v)^4 \tag{1}$$

where γ is the surface tension of a liquid, d_l and d_v are, respectively, the density of the liquid and vapor at a given temperature, and C is a constant characteristic of the substance in question. If the fourth root of both sides of equation (1) is taken, we obtain

$$\gamma^{1/4} = C^{1/4}(d_l - d_v)$$

and hence on solution for $C^{1/4}$ and multiplication of both sides by the molecular weight of the liquid, M, there follows

$$C^{1/4} = \frac{\gamma^{1/4}}{(d_l - d_v)}$$

$$P = MC^{1/4} = \frac{M\gamma^{1/4}}{(d_l - d_v)} \tag{2}$$

The quantity P defined by equation (2) in terms of the molecular weight, the surface tension, and the densities of the liquid and vapor is called the *parachor*. Since $P = MC^{1/4}$, and since both M and C are constants independent of temperature and characteristic for any given liquid, *the parachor for any substance should also be a constant independent of temperature and determined only by the nature of the liquid.* Of necessity d_l, d_v, and γ must all be determined at the same temperature for use in equation (2). When the temperature at which these are measured is considerably removed from the critical, d_v may be neglected in comparison to d_l, and we may write

$$P = \frac{M\gamma^{1/4}}{d_l} = \gamma^{1/4}V_m \tag{3}$$

where $V_m = M/d_l$ is the molar volume.

The physical significance of the parachor may be obtained by comparing the parachors of two substances. The ratio of the parachors of two substances A and B will be given by

$$\frac{P_A}{P_B} = \frac{\gamma_A^{1/4} V_{m(A)}}{\gamma_B^{1/4} V_{m(B)}}$$

At a temperature where the two liquids have equal surface tensions, $\gamma_A = \gamma_B$. Under such conditions, therefore, the ratio of the two parachors will be the ratio of the two molar volumes. Consequently the comparison of the parachors of various substances is essentially equivalent to a comparison of their molar volumes under conditions of equal surface tension, and hence internal pressure. This is the reason for the name parachor, which means comparative volume.

By comparing the measured parachors of various related substances Sugden was able to show that the parachor is both an additive and constitutive property of a liquid. The contributions to the parachor of a compound made by various atoms, bonds, and structures as thus found are summarized in Table 1. From this table it is immediately evident that not only the kind and number of atoms involved in a molecule determine the parachor, but also in certain instances their manner of arrangement and binding. Thus the presence of a double bond contributes 23.2 units, of a triple bond 46.6 units, while the difference between cyclic and noncyclic arrangements appears in the ring contributions.

<div align="center">TABLE 1</div>

<div align="center">ATOMIC AND STRUCTURAL PARACHOR CONTRIBUTIONS [1]</div>

Carbon	4.8	Double bond	23.2
Hydrogen	17.1	Triple bond	46.6
Nitrogen	12.5	3-membered ring	16.7
Oxygen	20.0	4-membered ring	11.6
O_2 in esters	60.0	5-membered ring	8.5
Chlorine	53.8	6-membered ring	6.1
Bromine	68.0	7-membered ring	4.6
Iodine	90.0	8-membered ring	2.4
Fluorine	25.0	Naphthalene ring	12.2
Sulfur	48.5	Semipolar bond	− 1.6
Phosphorus	39.2		

The use of Table 1 in calculating the parachors of various substances and in deducing from these the molecular structures may be illustrated with the following examples. Suppose it is desired to verify the structure of benzene, C_6H_6. The generally accepted structure of this compound is form I below,

[1] Glasstone, "Recent Advances in Physical Chemistry," D. Van Nostrand Company, New York, 1936.

$$
\begin{array}{ccc}
\overset{\displaystyle CH}{\underset{\displaystyle CH}{\underset{\displaystyle}{\parallel}}} & & \\
\end{array}
$$

$$
\begin{array}{ll}
HC-C\equiv C-C\equiv C-CH \\
\text{H} \qquad\qquad\qquad\qquad \text{H}
\end{array}
$$

I   II   III
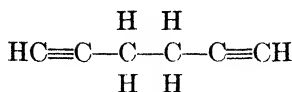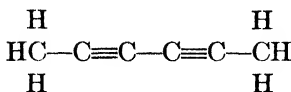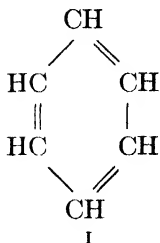
but forms II and III are also possibilities. Calculating the parachors for these three forms we have:

I	II
$6\ C = 6 \times 4.8 = 28.8$	$6\ C = 6 \times 4.8 = 28.8$
$6\ H = 6 \times 17.1 = 102.6$	$6\ H = 6 \times 17.1 = 102.6$
$3(\equiv) = 3 \times 23.2 = 69.6$	$2(\equiv) = 2 \times 46.6 = \underline{93.2}$
$1\ \bigcirc = 1 \times 6.1 = \underline{6.1}$	224.6
207.1	

III

$$
\begin{aligned}
6\ C &= 6 \times 4.8 = 28.8 \\
6\ H &= 6 \times 17.1 = 102.6 \\
2(\equiv) &= 2 \times 46.6 = \underline{93.2} \\
&\qquad\qquad\qquad 224.6
\end{aligned}
$$

The measured parachor of benzene is 206.3. This result is evidently in good accord with the calculated value for form I and not with the others. The accepted ring structure is thereby confirmed.

As another example we may take nitrobenzene, $C_6H_5NO_2$. For this compound we may write the two alternate structures

I   II
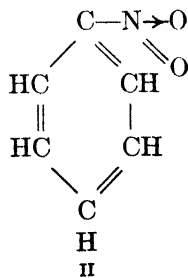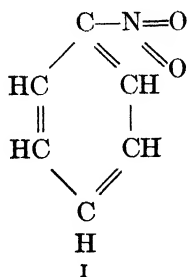
where in I the nitro group contains two oxygens doubly bonded to the nitrogen, while in II one is doubly bonded and one held by coordinate covalence ($N \rightarrow O$). For these two possibilities the calculated parachors are

	I			II	
6 C	=	28.8	6 C	=	28.8
5 H	=	85.5	5 H	=	85.5
1 N	=	12.5	1 N	=	12.5
2 O	=	40.0	2 O	=	40.0
1 ⬡	=	6.1	1 ⬡	=	6.1
5(=)	=	116.0	4(=)	=	92.8
		288.9			265.7

1 semipolar
bond (N → O) − 1.6
264.1

The experimental value 264.5 agrees excellently with structure II. Consequently in nitrobenzene, and in fact in other nitro compounds, one of the oxygens is doubly bound to the nitrogen, the second by coordinate covalence.

These examples illustrate some of the structural problems that may be solved with the aid of the parachor. The general procedure is first to measure the parachor of a compound by determining its surface tension and density, and next to compare this with the values calculated from Table 1 for various possible structures. When the calculated and measured parachors agree, the indications are that the given atomic arrangement is correct. When they do not, certain possibilities are eliminated. Used in this manner the parachor offers to the chemist a very simple and yet powerful tool for determining the structure of both organic and inorganic liquids. However, in one important instance the parachor is of little help, namely, in distinguishing structural isomers which do not differ from each other by a multiple bond or ring. For example, the measured parachors of o-, m-, and p-toluonitriles are respectively 290.6, 295.6, and 294.4, whereas calculation yields the same value for all, 292.9. Again, for a pair of isomers such as methyl valerate, $CH_3CH_2CH_2CH_2COOCH_3$, and propyl propionate, $CH_3CH_2COOCH_2CH_2CH_3$, both of which have the formula $C_6H_{12}O_2$, there is only one calculated parachor, 294.0, while two different values are observed, 292.5 for the first and 295.3 for the second. On the other hand, when two structural isomers differ from each other by a ring, as in the possibilities discussed for benzene, or by the number of double or triple bonds present, differentiation by means of the parachor is possible.

The Molar Refraction. Another quantity which is also an additive and partly constitutive property of a substance is the *molar refraction*. Lorenz and Lorentz showed in 1880 that on the basis of the electromagnetic theory of light the quantity

$$R_s = \left(\frac{n^2 - 1}{n^2 + 2}\right)\frac{1}{d} \tag{4}$$

should be a constant independent of temperature for any given substance. In this equation R_s is the *specific refraction* of the substance, n is the index of refraction, and d is the density measured at the same temperature as n. From equation (4) the molar refraction R_m is obtained on multiplying R_s by the molecular weight of the substance M, namely,

$$R_m = R_s M = \left(\frac{n^2 - 1}{n^2 + 2}\right)\frac{M}{d} \tag{5}$$

and R_m is also a constant.

The index of refraction of any medium is the ratio of the velocity of light in a vacuum to that in the given medium. It is measured by means of such optical instruments as the Pulfrich, Abbé, or immersion refractometers. Measured in this manner the index of refraction is not constant, but increases as the wave length of light is decreased; i.e., n is greater for violet light than for light toward the red end of the spectrum. To obtain comparable results it is necessary, therefore, to use light of a fixed frequency. For this purpose the common practice is to employ the yellow light of the sodium-D line from, say, a sodium vapor lamp. However, white light may be used in such instruments as the Abbé or immersion refractometers. In these instruments indices of refraction corresponding to sodium-D light can be obtained with ordinary light by the use of an incorporated prism which eliminates color fringes and reduces the results to sodium light. Further, in the Abbé refractometer the scale is calibrated directly in values of n, and hence the latter can be read immediately without any calculations.

The molar refraction R_m defined by equation (5) has been demonstrated to be an additive property for any given substance provided all measurements are referred to a given wave length of light. The atomic and various bond contributions evaluated for sodium-D light by inspection of observed data are given in Table 2. The molar refraction, like the parachor, may be seen to depend both on the number and nature of the atoms present, and also on the character of the binding. These values may be used to compare calculated with observed molar refractions in the same manner as explained for the parachor. Therefore the molar refraction, like the parachor, may be used to check the structure of molecules.

To illustrate the calculation of molar refractions we may take acetic acid, CH_3COOH. For this substance the density at 22.9° C is 1.046, the index of refraction at the same temperature for sodium light is 1.3715,

while the molecular weight is 60.05. The observed value of the molar refraction is, therefore,

$$
\begin{aligned}
R_m &= \left(\frac{n^2 - 1}{n^2 + 2}\right)\frac{M}{d} \\
&= \left[\frac{(1.3715)^2 - 1}{(1.3715)^2 + 2}\right]\frac{60.05}{1.046} \\
&= 13.03
\end{aligned}
$$

To compare with this value we find from Table 2 for acetic acid

$$
\begin{aligned}
2\,C &= 2 \times 2.418 = 4.836 \\
4\,H &= 4 \times 1.100 = 4.400 \\
1\,O- &= 1 \times 1.525 = 1.525 \\
1\,O= &= 1 \times 2.211 = \underline{2.211} \\
& 12.972
\end{aligned}
$$

The two values are seen to agree within 0.06 unit. Since the index of refraction is dimensionless, the units of the molar refraction are the units of M/d, i.e., volume. R_m is expressed thereby in *cubic centimeters*.

TABLE 2

MOLAR REFRACTION CONTRIBUTIONS
(For Sodium-D Light)

Carbon	2.418	Oxygen (in OH group, O—)	1.525
Hydrogen	1.100	Oxygen (in CO group, O=)	2.211
Chlorine	5.967	Oxygen (in ethers, O—)	1.643
Bromine	8.865	3-membered ring	0.71
Iodine	13.900	4-membered ring	0.48
Double bond	1.733		
Triple bond	2.398		

Calculated and observed molar refractions as a rule agree very closely. Exceptions arise only in open chain molecules with *conjugated* double bonds and in certain ring systems where the phenomenon of *optical exaltation* is observed; namely, that the observed value is generally higher than the calculated. Further, the molar refraction principle can be applied to gases, liquids, as well as solids. The molar refraction of gases is usually found to be within a few per cent identical with that of the corresponding liquids. Solids are studied best by dissolving them first in a solvent and measuring the refractive index n and density d of the solution. The molar refraction of the solution, $R_{1,2}$, follows then as

$$
R_{1,2} = \left(\frac{n^2 - 1}{n^2 + 2}\right)\left(\frac{N_1 M_1 + N_2 M_2}{d}\right) \tag{6}
$$

where M_1 and M_2 are the molecular weights of the solvent and solute respectively, while N_1 and N_2 are their mol fractions in solution. $R_{1,2}$

is in turn related to the individual molar refractions R_1 and R_2 by the equation

$$R_{1,2} = N_1 R_1 + N_2 R_2 \tag{7}$$

and hence when R_1 of the solvent is known, R_2 may be found from equation (7). This value of the molar refraction of the solute may be compared with values calculated from the molar refraction contributions given in Table 2. Such application of molar refraction to mixtures works out very satisfactorily.

Electrical Polarization of Molecules. Any molecule is composed of positively charged nuclei and negatively charged electrons. When such a molecule is introduced into an electric field between two charged plates, the field will act to attract the positive nuclei toward the negative plate and the electrons toward the positive plate, as shown in Fig. 1(a). The result is thus an electrical *distortion* or *polarization* of the molecule such that an *electric dipole* is formed, Fig. 1(b), with the positive charge at one end and the negative charge at the other. This polarization lasts only as long as the field is applied. When the field is removed the distortion disappears, and the molecule reverts to its original condition. For

Fig. 1. Polarization of a Molecule in an Electric Field; (a) Original State, (b) Polarized

this reason this type of electrical distortion of a molecule is called *induced polarization*, and the electric dipole formed is called an *induced dipole*.

Unlike the original molecule, where the centers of positive and negative electricity were coincident, in the induced dipole these charges are separated by a distance l. Again, since the molecule as a whole is neutral, the positive charge at one end Z_+, must be equal but be opposite in sign to the negative charge, Z_-, at the other end. To such an induced dipole may be ascribed an electric moment μ_i, which by definition is the value of the charge at *one* end of the dipole multiplied by the distance between the charges, namely,

$$\mu_i = Zl \tag{8}$$

The subscript i is used to denote that the dipole moment is induced only in an electric field.

The magnitude of the induced dipole moment μ_i is determined by the electric field strength X acting on the molecule in accordance with the relation

$$\mu_i = \alpha X \tag{9}$$

Here α is a constant called the *polarizability* of the given molecule. Clausius and Mosotti have shown from electromagnetic theory that the constant α is in turn related to the *dielectric constant* D of the medium between the plates, i.e., of the molecules, by the equation

$$\left(\frac{D-1}{D+2}\right)\frac{M}{d} = \frac{4}{3}\pi N\alpha = P_i \tag{10}$$

In this equation M and d are again the molecular weight and density of the molecules, while N is Avogadro's number. Since both N and α are constants independent of temperature, the quantity P_i, defined by equation (10), should also be a constant independent of the temperature and should be determined only by the nature of the molecules. P_i is called the *induced molar polarization* of the given substance, and it gives the electrical distortion produced in 1 mole of a substance when the electric field strength is unity. As D is unitless, the units of P_i are those of M/d, namely volume, and are expressed in cubic centimeters.

The dielectric constant is a property of any given medium. For a vacuum, $D = 1$, but for any other medium D is greater than unity. The dielectric constant of a substance can be determined by measuring first the capacity of a condenser with a vacuum between the plates, C_0, and next the capacity of the same condenser when filled with the given substance, C. Then the dielectric constant follows from electrical theory as

$$D = \frac{C}{C_0} \tag{11}$$

To measure the capacity, various electrical circuits are available, most of which employ vacuum tubes. All of them use alternating currents with frequencies of 10^6 to 10^7 cycles per second.

With the dielectric constant and the density of a substance measured at a given temperature, the molar polarization can readily be calculated by equation (10). When this is done it is found that in line with the above arguments the molar polarizations of such substances as oxygen, carbon dioxide, nitrogen, and methane are constant and independent of the temperature. On the other hand, for substances such as hydrogen chloride, chloroform, nitrobenzene, and methyl chloride the molar polarization is not constant but decreases as the temperature is raised. The explanation of this anomalous behavior by substances of the type mentioned has proved of great import to studies of molecular structure, and has yielded valuable information about molecules.

Permanent Dipole Moments. To account for the variation of the molar polarization of certain substances with temperature, Peter Debye made in 1912 the significant suggestion that molecules not only may develop an induced dipole moment in an electric field, but also may possess a *permanent* dipole moment of their own. The permanent dipole

moment would arise from the fact that in some molecules the centers of positive and negative electricity may not coincide. As a result there would be present in the molecule even outside of an electric field positive and negative centers of electrical gravity a distance l apart; and, consequently, such a molecule will be a permanent dipole and possess a permanent dipole moment μ of magnitude Zl. Outside of an electric field the permanent dipoles of various molecules in an aggregation would, because of thermal agitation, be oriented more or less randomly in space. However, when placed in an electric field two disturbing effects will arise. First, the electric field will tend to rotate and orient the permanent dipoles in the direction of the field; and second, the field will tend to polarize the molecules. Were the molecules perfectly stationary, the orienting effect of the electric field would result in an alignment of the dipoles at an angle of 180° to the direction of the field and 90° to the condenser plates. But the perfect alignment of the dipoles is disturbed by thermal agitation of the molecules, and hence the resulting orientation will be some point intermediate between the original position of the dipole in space and the final completely oriented position. This situation in an electric field is illustrated in Fig. 2, where (a) gives the initial orientation of the molecule with the permanent dipole, (b) the orientation resulting from the effect of the applied field.
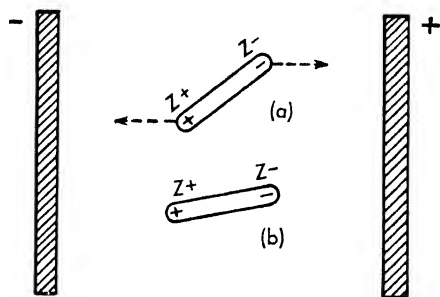
Fig. 2. Polarization of a Molecule with Permanent Dipole; (a) Original State, (b) Polarized and Oriented Position

In absence of a permanent dipole moment in a molecule the molar polarization, as calculated from the measured dielectric constant by means of equation (10), gives only the induced molar polarization P_i. On the other hand, when a dipole moment is present, this quantity measures not only the induced polarization, but also the *molar orientation polarization* P_o. The *total molar polarization* P_t is, therefore,

$$P_t = \left(\frac{D-1}{D+2}\right)\frac{M}{d} = P_i + P_o \qquad (12)$$

As before, the induced molar polarization is given by the Clausius-Mosotti equation, namely, $P_i = 4/3\,\pi N\alpha$. Again, Debye showed that the orientation polarization P_o should be equal to

$$P_o = \frac{4}{3}\pi N\left(\frac{\mu^2}{3\,kT}\right) \qquad (13)$$

where μ is the *permanent dipole moment* of the molecule, k is the gas constant per molecule, i.e., R/N, and T is the absolute temperature. Substituting these values of P_i and P_o into equation (12), the total molar polarization, as given by the measured dielectric constant, becomes

$$P_t = \left(\frac{D-1}{D+2}\right)\frac{M}{d} = \frac{4}{3}\pi N\alpha + \frac{4}{3}\pi N\left(\frac{\mu^2}{3\,kT}\right) \qquad (14)$$

In equation (14) the first term on the right is a constant which may be represented by A. In the second term, again, all quantities are constant except T, and hence this term may be written as B/T. With these substitutions, equation (14) takes the form

$$P_t = \left(\frac{D-1}{D+2}\right)\frac{M}{d} = A + \frac{B}{T} \qquad (15)$$

which shows that *for molecules with permanent dipoles the total molar polarization P_t should vary linearly with $1/T$.* Further, the slope of the plot of P_t vs. $1/T$ should be equal to

$$B = \frac{4\,\pi N\mu^2}{9\,k} \qquad (16a)$$

from which, on insertion of the values of N and k, μ should follow as

$$\mu = 0.0128\sqrt{B} \times 10^{-18} \qquad (16b)$$

Therefore, by plotting P_t against $1/T$ and taking the slope B of the linear plot, μ may be calculated from equation (16b), and thereby the permanent dipole moment of the molecule in question may be found.

Debye's theory of permanent dipole moments is confirmed by experimental data in every respect. In Fig. 3 are given plots of P_t vs. $1/T$ for hydrogen chloride and methyl chloride which show that for these substances P_t varies linearly with $1/T$, and hence these substances must possess permanent dipole moments. For comparison are included also the molar polarizations of methane and carbon tetrachloride at several temperatures. Since the molar polarizations of these substances are independent of tem-

Fig. 3. Variation of Total Polarization with Temperature for Polar and Nonpolar Compounds

perature, no permanent dipole moments are present and all electrical polarization is due, therefore, to induction.

Molecules possessing permanent dipole moments are said to be *polar*, while those without permanent dipoles are said to be *nonpolar*. The units of the dipole moment are those of the electronic charge multiplied by a distance. As the electronic charge is of the order of 10^{-10} electrostatic units and molecular distances are of the order of 10^{-8} cm, the dipole moments of molecules will always be of the order of magnitude of $10^{-10} \times 10^{-8} = 10^{-18}$. The quantity 10^{-18} is called a *Debye unit*, symbol D. A dipole moment may be given thus as, say, 1.6×10^{-18} or $1.6 D$.

Determination of Dipole Moments. The Debye equation (14) should be applicable primarily to gases and vapors. For such substances all that need be done is to measure the dielectric constants and densities at several temperatures, calculate from these P_t, and plot the latter against $1/T$. If the plot is independent of temperature, the substance possesses no permanent dipole moment. If, however, P_t varies with $1/T$, a permanent dipole is present and μ is evaluated from the slope by equation (16b).

With substances which are not gaseous, or which cannot be converted readily to a gas, an alternate procedure is employed. First several *dilute* solutions of the substance to be investigated in a *nonpolar* solvent, such as benzene, carbon tetrachloride, or carbon disulfide, are prepared, and the dielectric constants and densities of these are measured at several temperatures. Next at each temperature are calculated the molar polarizations of the solutions, $P_{1,2}$, from the relation

$$P_{1,2} = \left(\frac{D-1}{D+2}\right)\left(\frac{N_1 M_1 + N_2 M_2}{d}\right) \tag{17}$$

where N_1 and N_2 are the mol fractions of solvent and solute in the solution, M_1 and M_2 the respective molecular weights, and d the density of the solution. Third, from $P_{1,2}$, defined also as

$$P_{1,2} = N_1 P_1 + N_2 P_2 \tag{18}$$

where P_1 and P_2 are the molar polarizations of the solvent and solute, P_2 of the solute may be calculated from the known values of N_1, N_2, and P_1. If the values of P_2 thus determined are constant at each temperature for a number of concentrations, P_2 is the molar polarization of the pure solute. If P_2 is found to vary with concentration due to effect of the solvent, however, P_2 is plotted against N_2, and the curve is extrapolated to $N_2 = 0$ to yield P_2 freed of solvent influence. Finally the values of P_2 due to solute alone thus obtained at each temperature are

plotted against $1/T$, and μ is found from the slope in the same manner as for gases.

Other procedures are also available, but these need not be described here. Some results for the permanent dipole moments of various substances obtained either in the gas phase or in solution are shown in Table 3.

<div align="center">

TABLE 3

DIPOLE MOMENTS OF VARIOUS MOLECULES [1]
(In Debye Units)

</div>

Inorganic Molecules	μ	Organic Molecules	μ
H_2, Cl_2, Br_2, I_2, N_2	0	CH_4, C_2H_6, C_2H_4, C_2H_2	0
CO_2, CS_2, $SnCl_4$, SnI_4	0	CCl_4, CBr_4	0
HCl	1.03	C_6H_6, naphthalene, diphenyl	0
HBr	0.78	CH_3Cl	1.86
HI	0.38	RCl	2.04
H_2O	1.84	CH_3Br	1.78
H_2S	1.10	RBr	2.0
HCN	2.93	RCN	3.57
NH_3	1.46	ROH	1.65
SO_2	1.6	RNH_2	1.3
N_2O	0.14	CH_3COOH	1.4
CO	0.12	p-dichlorbenzene	0
PH_3	0.55	m-dichlorbenzene	1.48
PCl_3	0.9	o-dichlorbenzene	2.25
$SbCl_3$	3.75	p-chlornitrobenzene	2.55
$AsCl_3$	2.15	m-chlornitrobenzene	3.38
$AgClO_4$	4.7	o-chlornitrobenzene	4.25

Molecular Structure and Dipole Moments. The absence of a dipole moment in molecules such as hydrogen, chlorine, and nitrogen definitely indicates that the electron pairs binding the atoms together are situated equidistant between the constituents of the molecule. Were this not true a dipole moment would be present. Again, in linear molecules such as carbon dioxide and carbon disulfide, even if the electron pairs are not equidistant between the atoms, the electric moment on one side of the molecule balances the moment on the other side, and the net dipole moment is therefore zero. The same considerations apply also to such molecules as stannic chloride, carbon tetrachloride, and the paraffin hydrocarbons, both saturated and unsaturated. In aromatic molecules such as benzene, naphthalene, and diphenyl the absence of a net dipole moment can be accounted most readily on the basis of a planar structure for these. On the other hand, the presence of a dipole moment in the hydrogen halides and carbon monoxide definitely points

---

[1] R in this table is C_2H_5 or any higher straight-chain paraffinic group.

to a nonequidistant distribution of the electron pairs between the two atoms. In all probability this pair lies closer to the halogens and oxygen than to the hydrogen or carbon, and this leads to the formation of a dipole. In a similar manner may be explained the dipole moments appearing in the alkyl halides, nitriles, amines, and alcohols.

The fact that water, hydrogen sulfide, and sulfur dioxide have appreciable dipole moments definitely rules out linear structures for these molecules. If we postulate that the two hydrogens in water and hydrogen sulfide, and the two oxygens in sulfur dioxide, lie on the same side of the central atom, there will be no tendency for the moments to cancel each other, and an appreciable moment will result. Quantitative calculations indicate that this is the only way in which the dipole moment of these molecules can be accounted for. In a like manner the dipole moment of ammonia can be explained on the basis of a triangular pyramidal structure in which the nitrogen is situated at the apex of the pyramid, and the hydrogens lie at the other three corners.

It is of interest to observe that in p-dichlorbenzene, where the chlorines are at opposite ends of the benzene ring, the dipole moment is zero, and hence electrical symmetry is present. However, when substitution of the second chlorine takes place in a meta or ortho position, the symmetry is destroyed and a dipole appears. Furthermore, the dipole is larger the greater the asymmetry, as may be seen from a comparison of the μ values of m- and o-dichlorbenzenes. On the other hand, when the two substituents in the benzene ring are different, as in chlornitrobenzene, dipole moments are present in all three forms. Nevertheless the dipole moment of the para form is least, and increases successively with meta and ortho substitution.

Optical Rotation.[1] In Chapter IV attention was directed to the fact that crystals such as quartz, due to their asymmetry, possess the ability to rotate the plane of polarization of light. Since in this instance the asymmetry is associated with the crystal as a whole, the rotatory power persists only as long as the substance is solid, and disappears on melting or solution. On the other hand, when the rotation of the plane of polarization arises from asymmetry *within* the molecule, as in optically active organic compounds, optical activity does not disappear on change of phase, but is exhibited also by the substance when molten, dissolved in an appropriate solvent, or vaporized.

The extent of rotation of polarized light exhibited by a given optically active pure substance depends on the nature of the substance, the temperature, the thickness of substance through which the light passes, and on the wave length of the light. For any given wave length λ of light and temperature t, the angle α through which the plane of polarization is

[1] The student is advised to reread in this connection pp. 120–121 in Chapter IV.

rotated by a substance of density d and thickness l in *decimeters* is given by

$$\alpha_t^\lambda = [\alpha]_t^\lambda \, ld \tag{19}$$

In this equation the superscript refers to a fixed wave length of light, usually the sodium-D line, and the subscript to a given temperature, while $[\alpha]_t^\lambda$ is a proportionality constant called the *specific optical rotation* of the given substance. Since under any specified condition of λ and t $[\alpha]_t^\lambda$ is constant, this quantity may be used to designate the optical rotatory power of the substance.

Although equation (19) can be applied conveniently to optically active liquids, in dealing with solids the usual practice is to dissolve them first in an inactive solvent and then measure the angle of rotation α_t^λ. Under these conditions α_t^λ varies also with the concentration of the active substance in solution. If we let w_2 be the number of grams of active material per 100 cc of solution, then α_t^λ is given by

$$\alpha_t^\lambda = [\alpha]_t^\lambda \, l \, \frac{w_2}{100} \tag{20}$$

from which the specific rotation of the dissolved substance follows as

$$[\alpha]_t^\lambda = \frac{100 \, \alpha_t^\lambda}{l w_2} \tag{21}$$

For dilute solutions in any given solvent the specific rotation is generally constant, but this is not true in more concentrated solutions. Again, the specific rotation varies also with the nature of the solvent. As a rule solvents with high dipole moments give lower specific rotations than those with low dipole moments.

Instead of reporting specific rotations, *molar rotations* frequently are given. The molar rotation of a substance is merely the specific rotation multiplied by the molecular weight, i.e.,

$$[M\alpha]_t^\lambda = [\alpha]_t^\lambda \, M. \tag{22}$$

In organic compounds optical activity is associated most often with molecules containing an asymmetric carbon atom, i.e., a carbon atom to which four *different groups* are attached. With such an atom present, two geometric configurations of the molecule are possible, namely,

<pre>
      A   A
      | |
   B—C—E   E—C—B
      | |
      D   D
      I   II
</pre>
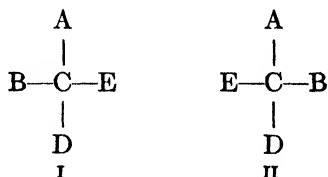
of which one is a minor image of the other. These two molecules are identical in all chemical and physical properties, except that one form will rotate the plane of polarized light to the right, while the other *exactly the same amount* to the left. When prepared *synthetically*, compounds containing an asymmetric carbon atom are always found to be optically inactive because of presence of the two active forms in exactly equivalent quantities. Such combinations of the two forms in equivalent amounts are called *dl-mixtures.* By various methods the dl-mixtures may be separated into the dextro and laevo forms.

Optical activity measurements have been applied not only to differentiate the active forms of various substances, but also to the study and analysis of sugars. By means of their optical rotation various sugars may be identified, and the concentrations of mixtures of various sugars may be ascertained with the aid of polarimeters.

Absorption of Light by Liquids and Solutions. A gaseous molecule on being irradiated with light will emit a band spectrum which may result from electronic, vibrational, or rotational transitions. In such an instance the molecule is first excited, and the radiations resulting on return of the molecule to its original state are observed. Investigation of the band spectrum of a molecule in this manner is called study of the molecule in *emission.* However, a molecule may also be studied in *absorption,* in which case the energies required for excitation of various transitions are ascertained. For this purpose a gaseous substance is exposed to continuous radiation from which the molecule removes light of frequencies which correspond to those necessary to cause rotational, vibrational, or electronic transitions. The frequencies thus absorbed by a molecule can readily be determined by inspecting the transmitted light in a spectrograph. There in place of a continuous spectrum will be found light portions interspersed with dark spaces which correspond to the frequencies of the removed light. Of necessity the positions of the dark spaces will be the same as those of emitted light for the same internal transitions.

Whereas a gas on absorption exhibits electronic bands with discernible vibrational structure, a liquid or solution yields as a rule absorption bands which are continuous. This continuity in the liquid state may be due either to overlapping of vibrational bands, or to destruction of quantized absorption by interaction of intermolecular with intramolecular energies. Whatever the case may be, the absorption appears usually as a continuous band extending over part of the spectrum. The spectral range where the maximum absorption takes place depends on the energy involved in the electronic transition responsible for the absorption. If the energy of the latter is large, the absorption will occur primarily in the ultraviolet. With smaller energies, on the other hand,

the absorption band will appear in the near ultraviolet or in the visible light regions.

The color exhibited by a substance is determined by the light the substance transmits in the visible range, and this, in turn, is related to the spectral range in which the substance absorbs light. Any substance that absorbs only in the ultraviolet will transmit all white light, and as a consequence the substance will appear colorless. If, however, the substance shows a color in white light, the color indicates the light which was *not* absorbed. For instance, carbon black has the particular color because it absorbs nonselectively the white light that reaches it. Again, sodium chloride solutions are colorless because they do not absorb selectively any light in the visible region. On the other hand, a copper sulfate solution is blue because it absorbs out of the white light the yellow and red, leaving the blue to be transmitted to the eye. It is evident that the characteristic of the absorbing substance is not the light it transmits but rather the light it absorbs. Consequently it is by studying the latter that information about the absorbing substance can be obtained. Further, it also follows that only colored substances can be studied in the visible region. Substances which are colorless do not absorb white light appreciably, and hence these must be investigated in the ultraviolet.

Spectrophotometry. In the spectral range where the absorption is appreciable the intensity of the initial light is of necessity reduced greatly on passing through the absorbing medium. On the contrary, to either side of the absorption band the intensity is reduced considerably less or not at all. From these considerations it is evident that the position of an absorption band can be ascertained by measuring the decrease in intensity of the original light after passage through an absorbing medium as a function of the wave length. Then, on plotting the fraction of the original light absorbed against the wave length, the position of the absorption band will be given by a maximum curve. Conversely, if we plot instead the fraction of light transmitted vs. the wave length, a minimum in the plot will indicate the position of maximum absorption.

The manner in which the intensity of incident light of *any given wave length* is reduced on passing through any pure absorbing substance is given by *Lambert's law*. This law states that the rate of decrease of intensity with thickness of absorbing material $-dI_t/dl$, where I_t is the intensity at thickness l, is proportional to the intensity of the light at any given point l; namely, that

$$-\frac{dI_t}{dl} = kI_t \tag{23}$$

In equation (23) the proportionality constant k is called the *absorption coefficient* and is characteristic of the absorbing medium. Since at

$l = 0$ we have the original intensity I_0, the intensity I_t at any point $l = l$ can be found from equation (23) by integration between these limits. We obtain thus

$$\int_{I_0}^{I_t} \frac{dI_t}{I_t} = \int_{l=0}^{l=l} - kdl$$

$$\ln \frac{I_t}{I_0} = - kl \tag{24a}$$

$$\frac{I_t}{I_0} = e^{-kl} \tag{24b}$$

The last equation may be written also as

$$\frac{I_t}{I_0} = 10^{-(2.303k)l}$$

$$= 10^{-k'l} \tag{25}$$

in which case $k' = 2.303\, k$ is referred to as the *extinction coefficient* of the substance. From equation (24a) it is evident that $\ln I_t$ of the transmitted light falls off linearly with the thickness traversed. Or, stated differently, the intensity of the transmitted light falls off exponentially with the distance l light has to travel through the absorbing medium.

For absorbing solutes dissolved in a solvent the decrease in intensity with l is proportional not only to I_t, but also to the concentration of the solution C. For solutions, therefore,

$$\frac{- dI_t}{dl} = \epsilon I_t C \tag{26}$$

where ϵ, called the *molar absorption coefficient*, is a proportionality constant determined by the nature of the absorbing solute and the wave length of light used. Integration of equation (26) between the same limits as before, and at constant C, yields

$$\int_{I_0}^{I_t} \frac{dI_t}{I_t} = -\int_{l=0}^{l=l} \epsilon C dl$$

$$\ln \frac{I_t}{I_0} = - \epsilon C l \tag{27a}$$

and

$$\frac{I_t}{I_0} = e^{-\epsilon C l} \tag{27b}$$

Equations (26) and (27) are expressions of *Beer's law* for absorption of light by solutions. These show that the intensity of the transmitted light falls off exponentially with both C and l, or that the logarithm of the transmitted intensity I_t is proportional to both the concentration and the length of the absorbing path. Equation (27a) indicates that

the same change in intensity can be produced by keeping l constant and varying C as by holding C constant and changing l proportionately.

As before, the base of natural logarithms can be changed in equation (27b) to base 10, in which case

$$\frac{I_t}{I_0} = 10^{-(2.303\varepsilon)Cl}$$
$$= 10^{-\varepsilon'Cl} \tag{28}$$

The constant $\varepsilon' = 2.303\ \varepsilon$ is called the *molar extinction coefficient* of the absorbing solution.

The ratio I_t/I_0 is determined with some form of spectrophotometer. At present spectrophotometers utilizing photoelectric cells are available which give I_t/I_0 directly. In these the instrument is first adjusted to read 100 per cent transmission when the pure solvent alone is interposed between the light source and the detecting photoelectric cell. The pure solvent is then removed, the solution is substituted, and the instrument is again balanced. From a dial suitably calibrated the ratio I_t/I_0 can be read immediately as per cent transmission. By repeating the procedure with various incident wave lengths of light, a complete transmission curve can be traced quite rapidly and conveniently.

Fig. 4. Absorption Curve for Chromium-Phenylcarbazide. (Reproduced by permission of Coleman Electric Co., Maywood, Ill.)

Once I_t/I_0 for a pure liquid is determined and the thickness of the cell used is known, the absorption or extinction coefficients of the substance for the given wave length can readily be calculated from equations (24b) or (25). Similarly from the measured I_t/I_0 and the known values of l for the cell, and C for the solution used may be found the molar absorption and extinction coefficients through equations (27b) and (28). Next, to ascertain the wave length at which maximum absorption takes place, we may plot (a) these coefficients or their logarithms against the wave length λ, or (b) the ratios I_t/I_0 or log I_t/I_0 vs. λ. On any one of these plots the wave length at which appreciable absorption occurs will be indicated by a minimum. An illustrative graph of per cent transmittance, i.e., $I_t/I_0 \times 100$, vs. λ, for the complex formed between chromium and phenylcarbazide in aqueous solution is shown in Fig. 4. Here an absorption band is clearly indicated at a wave length

of 5400 Å. Although in this instance only one absorption band appears in the visible range, in other substances two and more such minima may be found. Each of these corresponds to a separate electronic excitation of the molecule. Other absorption bands may also be detected in the ultraviolet.

Absorption spectrophotometry finds application in the definition of the spectral colors of indicators, in the exact definition of color of solutions, liquids, and solids, in analysis, and in the solution of various structural problems. Reference to the use of spectrophotometry in analysis will be made below. At present a classical example of absorption study to the solution of a chemical structural uncertainty will be given. For the compound isatin two structural formulas were proposed, namely,

$$(a)\ C_6H_4 \begin{array}{c} CO \\ \diagup\ \diagdown \\ \diagdown\ \diagup\!\!\!= \\ N \end{array} C\text{—}OH \quad \text{and} \quad (b)\ C_6H_4 \begin{array}{c} CO \\ \diagup\ \diagdown \\ \diagdown\ \diagup \\ NH \end{array} CO$$
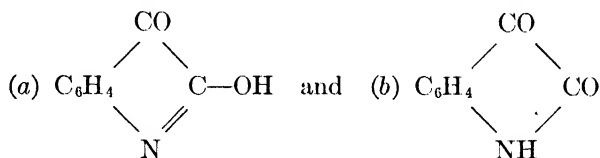
The question arose, then, which of these is the correct one. Now, from isatin two ethers can be prepared whose structural formulas are

$$C_6H_4 \begin{array}{c} CO \\ \diagup\ \diagdown \\ \diagdown\ \diagup\!\!\!= \\ N \end{array} C\text{—}OC_2H_5 \qquad C_6H_4 \begin{array}{c} CO \\ \diagup\ \diagdown \\ \diagdown\ \diagup \\ N\text{—}C_2H_5 \end{array} CO$$

Lactim ether   Lactam ether
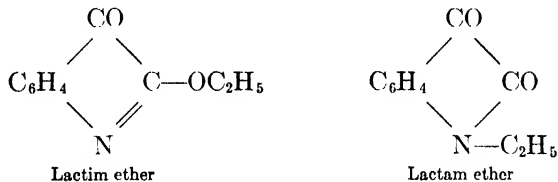
To settle the problem of the structure of isatin Hartley and Dobbie (1900) studied the absorption spectra of the two ethers and of isatin itself. They found that the absorption curve of the lactam ether was very similar to that of isatin, but that of the lactim ether was quite different. They concluded, therefore, that the structure of isatin is similar to that of the lactam ether, and hence that this substance has the formula (b). In an analogous manner absorption spectrophotometry has been applied to other compounds, both organic and inorganic, and has proved very useful in shedding light on their structure.

Spectrophotometric Analysis. Spectrophotometry is used extensively at present for analytical purposes. A solute to be suitable for photometric estimation in the visible range must exhibit absorption in this range, i.e., it must have a color in solution. If the solution is not colored, or if the color is not satisfactory, two possibilities are open. First, ultraviolet instead of visible light may be employed, or, second, some substance may be added which reacts with the solute to yield a

colored product. The latter procedure is exemplified by analysis of ferrous iron solutions by addition to these of phenanthroline. The phenanthroline reacts with ferrous ions to form a red-colored complex, the intensity of whose color in solution is proportional to the concentration of the ferrous ions. Consequently, by determining the absorption variation of various ferrous-phenanthroline complex solutions with the concentration of ferrous ions, the latter may be estimated.

In absorption studies the cell thickness and solution concentration are kept constant, and I_t/I_0 at various wave lengths is measured. On the other hand, in analysis I_t/I_0 is desired not as a function of λ but of the concentration C. To obtain this relation a cell of fixed thickness is used, and the wave length of the light is kept constant at some value near the absorption maximum of the substance to be analyzed. The first task in any spectrophotometric estimation is the preparation of a calibration curve, i.e., a plot of $\log I_t/I_0$ vs. concentration. For this purpose several solutions of known concentration of the substance to be analyzed are prepared, their intensity ratios are read with the spectrophotometer, and a plot of $\log I_t/I_0$ vs. C is constructed. Next the unknown solution is inserted in the instrument, and the intensity ratio for it is measured. Knowing thus $\log I_t/I_0$ for the unknown, the concentration corresponding to this value of the ratio is read off the curve, and this is the concentration of the given solution.

In Fig. 5 is shown a typical $\log I_t/I_0$ vs. C calibration plot, in this instance for the estimation of chromium by the phenylcarbazide method. The fact that the line is straight is an indication that these solutions obey Beer's law. When Beer's law is not followed, as is frequently the case when white instead of monochromatic light is used, the line is not straight. However, whether Beer's law is obeyed or not is immaterial here. The calibration curve may be used in any case to obtain the concentration from the measured intensities.
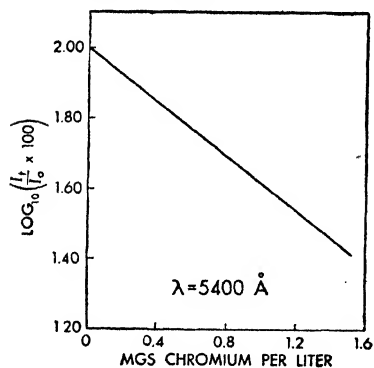
Fig. 5. Spectrophotometer Calibration for Chromium Phenylcarbazide Method. (Reproduced by permission of Coleman Electric Co., Maywood, Ill.)

One advantage of spectrophotometric methods over other schemes of analysis is that they permit determination of concentrations of very dilute solutions. Thus Fig. 5 shows that as little as 0.1 mg of chromium per liter of solution may be estimated. Similarly phosphate solutions as dilute as 10^{-6} or 10^{-7} molar may be analyzed spectrophotometrically.

In fact, the absorption of more concentrated solutions is generally so great that for accuracy it is desirable to dilute these sufficiently to yield per cent transmittances somewhere between 20 and 80.

REFERENCES FOR FURTHER READING

1. Baly, "Spectroscopy," Longmans, Green, and Company, London, 1927.
2. W. R. Brode, "Chemical Spectroscopy," John Wiley & Sons, Inc., New York, 1943.
3. P. Debye, "Polar Molecules," Reinhold Publishing Corporation, New York, 1929.
4. S. Glasstone, "Textbook of Physical Chemistry," D. Van Nostrand Company, Inc., New York, 1940.
5. S. Glasstone, "Recent Advances in Physical Chemistry," The Blakiston Company, Philadelphia, 1936.
6. R. J. W. Le Fèvre, "Dipole Moments," Chemical Publishing Company of New York, Inc., New York, 1938.
7. T. M. Lowry, "Optical Rotatory Power," Longmans, Green and Company, London, 1935.
8. T. M. Lowry, "Optical Rotation of Liquids," Miscellaneous Publication No. 118, U. S. Bureau of Standards, Washington, D. C., 1932.
9. C. P. Smyth, "Dielectric Constant and Molecular Structure," Reinhold Publishing Corporation, New York, 1931.
10. S. Sugden, "The Parachor and Valency," George Routledge & Sons, Ltd., London, 1930.

PROBLEMS

1. Predict from calculated parachor values the surface tensions at 20° C of cyclohexane and brombenzene, and compare with observed values. The densities of the two substances are, respectively, 0.779 and 1.495 g/cubic centimeter.

   Ans. Calc. 24.5 and 36.4; obs. 25.3 and 36.5 dynes/cm

2. Calculate from Table 1 the parachor values of CH_3COOH and benzonitrile (C_6H_5CN), and compare them with the values calculated from the surface tensions and densities of these compounds.

3. The surface tension of allyl alcohol is 25.8 dynes/centimeter and the density 0.854 g/cubic centimeter at 20° C. From parachor evidence which of the following two structures should be correct

$$CH_2{=}CH{-}CH_2OH \quad \text{or} \quad \begin{matrix} CH_2 \\ | \\ CH_2 \end{matrix}\!\!\!\Big\rangle CHOH$$
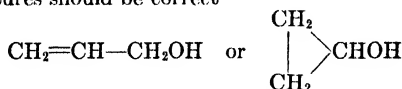
4. The density of allyl chloride ($CH_2{=}CH{-}CH_2Cl$) is 0.938 g/cubic centimeter at 20° C. Calculate the index of refraction, and compare your answer with the observed value of 1.4154.

5. The refractive index of naphthalene decahydride ($C_{10}H_{18}$) is 1.4804 at 18° C. Calculate the density and compare with the observed value of 0.895.

6. Calculate the molar refractions of the following compounds:

Compounds	Density	Refractive Index
Methyl chloroformate	1.223	1.3868
Tertiary butyl chloride	0.843	1.3869
Secondary butyl alcohol	0.806	1.3924

and compare your results with those obtained from the additivity rule.

7. The density of ethyl alcohol at 20° C is 0.789 g/cubic centimeter while that of methyl alcohol is 0.792 g/cubic centimeter. Assuming that these substances form an ideal solution, calculate by using Table 2 the refractive index of a solution containing 50% by weight of each constituent.   *Ans.* 1.2353

8. From the following data giving the dielectric constant D at various temperatures, determine the dipole moment of gaseous HCl by a graphical method:

D	$t°$ C
1.0076	− 75
1.0046	0
1.0026	100
1.0016	200

The pressure is 1 atm in every case. Assume the ideal gas law in calculating the density of HCl.   *Ans.* 1.13 D

9. From the dipole moment calculate the mean distance between the center of positive and negative electricity in the HCl molecule.

10. The density of *d*-butanol-2 is 0.808 g/cubic centimeter at 20° C, and the specific rotation, $[\alpha]_{20°}^{D}$, is 13.87°. What should be the angle of rotation, $\alpha_{20°}^{D}$, of a sample of butanol-2 contained in a 5 cm tube? What is the molecular rotation?   *Ans.* 5.50°

11. The specific rotation of *d*-butanol-2 in CS_2, $[\alpha]_{20°}^{5461}$, is 20.46°. Calculate the angle of rotation, α_{20}^{5461}, of a 0.1 molar solution of *d*-butanol-2 in CS_2 measured in a 10 cm tube.

12. The optical rotation of a solution containing 5 g *d*-sec butyl acetate in 100 cc of alcohol, α_{20}^{D}, is 1.29° in a tube 10 cm long. Calculate the specific rotation of *d*-sec butyl acetate, $[\alpha]_{20}^{D}$, and the molecular rotation.

13. The following transmission data for K_2CrO_4 solutions were taken at a wave length of 3660 Å with a photoelectric spectrophotometer:

I_t/I_0	Concentration × 10^4 moles/liter
0.42	0.8
0.275	1.2
0.175	1.6
0.110	2.0

From these data test Beer's law graphically, and from your plot determine the molar extinction and the molar absorption coefficients of K_2CrO_4 solutions at this wave length.

Index